Large-Scale and Linear-Scaling Quantum Mechanics Computational Methods to characterize the DNA G-quadruplexes and their interaction with small molecules

Eskala-handiko eta Eskalatze–Linealeko Mekanika Kuantikozko Simulazio Konputazionalak ADN G-quadruplexuak eta molekula txikiekin dituzten interakzioak ezaugarritzeko

Iker Ortiz de Luzuriaga López

2023

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea
NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

IKER ORTIZ DE LUZURIAGA

LARGE-SCALE AND LINEAR-SCALING QUANTUM MECHANICS COMPUTATIONAL METHODS TO CHARACTERIZE THE DNA G-QUADRUPLEXES AND THEIR INTERACTION WITH SMALL MOLECULES

LARGE-SCALE AND LINEAR-SCALING QUANTUM MECHANICS COMPUTATIONAL METHODS TO CHARACTERIZE THE DNA G-QUADRUPLEXES AND THEIR INTERACTION WITH SMALL MOLECULES

IKER ORTIZ DE LUZURIAGA

CIC nanoGUNE
MEMBER OF BASQUE RESEARCH & TECHNOLOGY ALLIANCE

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

DIPC

Director:
Adrià Gil

Co-director:
Xabier López

Iker Ortiz de Luzuriaga: *Large-Scale and Linear-scaling Quantum Mechanics Computational Methods to Characterize the DNA G-quadruplexes and their interaction with small molecules*

Amarentzat

ABSTRACT

G-quadruplexes have raised considerable interest in developing therapies against cancer during the last few years. These non-canonical structures of DNA may be found in telomeres and/or oncogene promoters, and it has been observed that the stabilization of such G-quadruplexes may disturb tumor cell growth. Nevertheless, the mechanisms leading to the folding and stabilization of these G-quadruplexes are still not well established, and they are the focus of current work in this field. In this thesis, the interaction of two isomers, equatorial and axial, of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex with different DNA structures was studied, taking particular interest in DNA G-quadruplexes. We use computational methods to gain insight into the experimentally found cytotoxicity. Among all the methods used, the Linear-Scaling Density Functional Theory has a particular relevance throughout the work. Still, techniques such as semi-empirical, DLPNO-CCSD(T), and QM/MM have also been used. Besides, we supplement the theoretical work employing EDA, QTAIM, and NCI analysis to get insight into the weak non-covalent interactions of these systems that modulate their affinity. Computed formation energies, energy decomposition analysis, solvation energies, and non-covalent interaction analysis explains the preference of the metal complex for the G-quadruplex DNA binding over the duplex DNA. We also observed that an axial complex is more favorable for interaction with the G-quadruplex DNA than the equatorial one. This is due to the conformation adopted by the axial complex, wholly inserted in the cavity of the G-quadruplex structure, disposed between tetrads, and favoring the establishment of stabilizing non-covalent weak interactions. The most relevant weak interactions correspond to $\pi - \pi$ stacking ones because the phenanthroline ligand's flat aromatic surface is perfect for interacting with the tetrads of the G-quadruplex DNA structure. On the other hand, we also found that the role of the ancillary ligands is crucial to enhance the interaction of the metal complex with DNA.

LABURPENA

G-quadruplexuek interes handia piztu dute azken urteetan minbiziaren kontrako terapiak garatzeko. DNAren egitura ez-kanoniko hauek telomeroetan eta/edo onkogenoen sustatzaileetan aurkitu daitezke, eta frogatu da egitura horien egonkortzeak tumore-zelulen hedapena eragotz dezakeela. Hala eta guztiz ere, G-quadruplexu egituren tolestura eta egonkortzea eragiten duten mekanismoak oraindik ez daude nahi bezain ondo finkatuak, eta gaur egungo lan askoren ardatza dira. Tesi honetan esperimentalki aurkitutako zitotoxikotasunaren jatorria ikertzeko DNA egitura ezberdinak ikertu dira, bereziki G-quadruplexua, eta hauen elkarrekintza [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] konplexu metalikoaren bi isomerorekin, ekuatoriala eta axiala. Ikerketa aurrera eramateko metodo ezberdinak erabili dira. Erabilitako metodoen artean, eskalatu-linealeko dentsitate funtzionalaren teoria garrantzia berezia du lan osoan zehar, baina eskalatu-lineal dentsitate funtzional theoriak garrantzi berezia du lan osoan zehar, baina erdi-enpirikoa, DLPNO-CCSD(T), QM/MM. Beste aldetik EDA, QTAIM eta NCI bezalako tresna konputazionalak ere erabili dira. Kalkulatutako eraketa-energiak, EDA analisiak, disolbatze-energiak eta interakzio ez-kobalenteen analisiak konplexu metalikoak duplex DNArekin baino, G-quadruplexuarekin elkarrekintza osatzeko lehentasuna duela frogatzen da. Bi isomeroen artean, konplexu axiala G-quadruplexuarekin elkarrekintza sortzeko erraztasun gehiago duela frogatu da, G-quadruplexu egituraren barnealdean guztiz txertatuta azaltzen delako, laukoteen artean jarrita, egonkortasuna areagotzen duten elkarrekintza ahulen formakuntza sustatuz. Elkarrekintza ahulen artean, garrantzitsuena $\pi - \pi$ elkarrekintza da, bere gainazal aromatikoa ezin hobea baita laukoteekin elkarreragiteko. Bestalde, gainontzeko ligandoen eginkizuna funtsezkoa da konplexu metalikoaren elkarrekintza aproposa lortzeko.

PUBLICATIONS

Some ideas and figures have appeared previously in the following articles, published during my PhD:

- Iker Ortiz de Luzuriaga, Xabier López, and Adrià Gil. "Learning to Model G-Quadruplexes: Current Methods and Perspectives." In: *Annual Review of Biophysics* 50.1 (2021), 209–243. doi: 10.1146/annurev-biophys-060320-091827
- Iker Ortiz de Luzuriaga, Sawssen Elleuchi, Khaled Jarraya, Emilio Artacho, Xabier López, and Adrià Gil. "Semi-empirical and Linear-Scaling DFT Methods to Characterize duplex DNA and G-quadruplexes in Presence of Interacting Small Molecules." In: *Physical Chemistry Chemical Physics*, 24 (2022), 11510-11519. doi: 10.1039/D2CP00214K
- Iker Ortiz de Luzuriaga, Ángel Sánchez-González, Wojciech Synoradzki, Xabier López, and Adrià Gil. "Unravelling the Binding Affinity and Selectivity of Molybdenum (II) Phenanthroline Complexes with DNA G-Quadruplexes by Using Linear-Scaling DFT studies. The Important Role of Ancillary Ligands.(2022)" In: *Phys.Chem.Chem.Phys.*, 24 (2022), 25918-25929. doi: 10.1039/D2CP02241A.
- Sawssen Elleuchi, Iker Ortiz de Luzuriaga, Ángel Sanchez-Gonzalez, Xabier López, Khaled Jarraya, Maria José Calhorda, and Adrià Gil. "Computational Studies on the Binding Preferences of Molybdenum(II) Phenanthroline Complexes with Duplex DNA. The Important Role of the Ancillary Ligands." In: *Inorganic Chemistry* 59.17 (2020), 12711–12721. doi: 10.1021/acs.inorgchem.0c01793

- Ángel Sánchez-González, Nuno A. G. Bandeira, Iker Ortiz de Luzuriaga, Frederico F. Martins, Sawssen Elleuchi, Khaled Jarraya, Jose Lanuza, Xabier López, Maria José Calhorda, and Adrià Gil. "New Insights on the Interaction of Phenanthroline Based Ligands and Metal Complexes and Polyoxometalates with Duplex DNA and G-Quadruplexes." In: *Molecules* 26.16 (2021), 4737. doi: 10.3390/molecules26164737

Other Publications:

- Elisa Jimenez-Izal, Iker Ortiz de Luzuriaga, Eloy Ramos-Cordoba, and Jon M. Matxain. "Role of Dispersion Interactions in Endohedral TM@(ZnS)12 Structures." In: *ACS Omega* 6.25 (2021), 16612–16622. doi: 10.1021/acsomega.1c02016

ESKERRIK ASKO

Azkenean, abentura honetan zehar inguratu nauen jendeari eskerrak emateko garaia iritsi zait. Hitz hauek idazterakoan sentsazio arraroa datorkit, atzo hasi nintzela iruditzen zaidalako, baina ¡Zenbat gauza pasa diren lau urte hauetan!

Primero de todo, quiero dar las gracias a Adrià y a Xabi por haber confiado en mi, y aunque se que en algunos momentos no ha sido fácil, siempre habéis estado apoyándome. No sabéis lo mucho que ha significado para mi.

Unibertsitateko jendeari. Txantxot, Maria, Maialen, Lanuza, Xabi Telle, David, Julen, Oksana, Andoni, Zubeltzu. Zer egingo nuke ni zuekin goizero kafea artuko ez banu. Eskerrik asko bihotz-bihotzez. Eta nola ez, Txoni eta Txema. Zuekin hasi nintzen ibilaldi honetan eta ¡A zer nolako nahaspilan sartu nintzen!

I am really grateful to all the nanopeople. There are too many of you to name all of you one by one, but writing these words I have you all in mind. You welcomed me with open heart and it has been four wonderful years by your side. All the memories that I take from your company are a present that I did not count on when I started this thesis. I will miss you.

Kuadrilla!! Honekin guztiarekin zerikusirik izan ez baduzue ere, zuekin partekatzen ditut barrerik eta momenturik alaienak. Izandako momentu atsegin guztiengatik, hiztegiko hitz guztiekin ere ezin izango nizueke nahi bezain beste zuen adiskidetasuna eskertu.

Garrantzitsuenak amaierarako utzi nahi izan ditut. Ama. Aita. Zuek zarete egunez egun jasan nauzuena, horrek dakarren guztiarekin. Eta nola ez, Julen, etxeko txikiari. Ez zaitez aldatu eta jarraitu beti bezain txoriburu eta burugogor. Abentura hau hemen amaitzen den arren, ziur nago nire ondoan egongo zaretela hemendik aurrera datozen hanka-sartze guztietan. Baina, mesedez eskatzen dizut ama, ez iezadazu berriro galdetu ¿Ba al dakizu hemendik aurrera zer egingo duzun? Erantzuna ez zaizu gustatuko.

Eskerrik asko denoi!!

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

BCP Bond Critical Point

CCSD Coupled-Cluster with iterative inclusion of single and double excitations

CCSD(T) Coupled-Cluster with iterative inclusion of single and double excitations with perturbative inclusion of triple excitations

CP Critical Point

CCP Cage Critical Point

QM/MM Quantum Mechanics/Molecular Mechanics

dDNA duplex DNA

DFT Density Functional Theory

DFT-D Density Functional Theory with dispersion corrections

DLPNO Domain-based Local Pair Natural Orbital

EDA Energy Decomposition Analysis

FF Force Fields

GQ G-quadruplex DNA

LS-DFT Linear-Scaling Density Functional Theory

MD Molecular Dynamics

mg minor groove

MG Major Groove

MM Molecular Mechanics

NCI Non-Covalent Interactions

NCP Nuclear Critical Point

phen phenanthroline

QM Quantum Mechanics

QTAIM Quantum Theory of Atoms In Molecules

RCP Ring Critical Point

SE Semi-Empirical

1

INTRODUCTION

Therapies to fight against tumoral cells is one of the most widely studied field by chemists and biologists. Worthwhile studies focus on drugs, which based their cytotoxic response on interacting with and altering the nuclear DNA, thereby interfering with the replication and transcription machinery of the cell [16, 28]. One of the best-known drug interacting with DNA is the cisplatin, since its introduction in chemotherapy treatments for cancer [96]. However, such drug does not distinguish between tumor and normal cells, and hence serious side effects occur due to damage of normal tissues [103, 137]. The toxicity of cisplatin makes that nowadays DNA platination is the ultimate event in the cancer therapy and opens the doors to new compounds. An alternative way of fighting cancer can be the use of non-canonical DNA structures such as G-quadruplex DNA (GQ) as molecular specific targets. These new targets encourage researchers to explore new ways to replace cisplatin by a new drugs.

1.1 G-QUADRUPLEX DNA

GQ's are non-canonical secondary structures formed from nucleic acid sequences that are rich in guanine. The basic unit of these structures is known as a G-quartet or G-tetrad and it consists of a square planar structure formed by four guanine bases, which interacts via Hoogsteen hydrogen bonds (Figure 1) [219]. Stacking of more than one G-tetrad forms the GQ structure (Figure 2). This stacking is produced in such a way that the O6 atoms of guanines are oriented to the midpoint of the structure, and the resulting system has some tubular empty space in

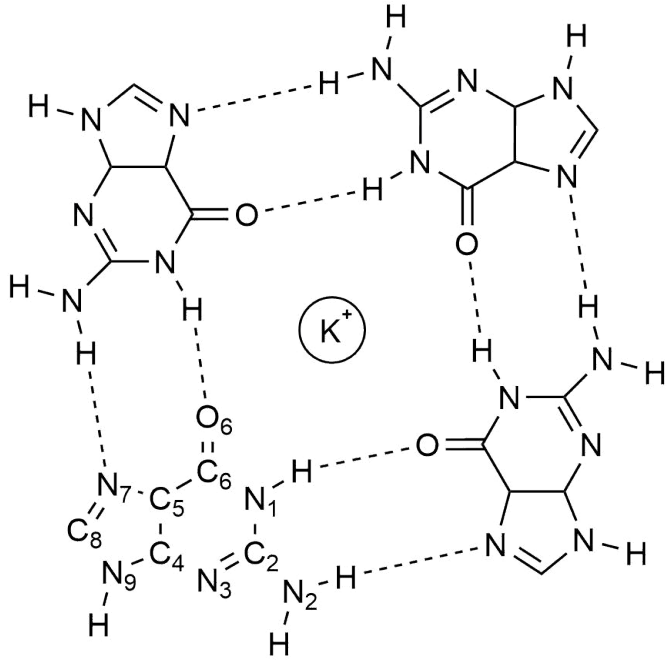

Figure 1: G-tetrad structure with a potassium cation.

the center, which is called the ion channel. When an alkali metal cation is placed in this ion channel [109], it may interact with up to eight O6 atoms; this coordination leads to the stabilization of the GQ. Moreover, from the guanines involved in the G-tetrads of the GQ's, various numbers of interspersing residues form loops that connect each G-tetrad plane (Figure 2). GQ's can be viewed and studied as monomeric structures that differ in tetrad number, strand polarity, loop topology, etc., or they can be studied as multimers that form higher-order structures, which can contain hundreds of GQ monomers [31, 132]. To form the GQ's, the regular dDNA structure must open; once the helix is divided, the guanine-rich strand forms the GQ's. Although GQ's coming from a single strand are the most common structures, GQ's derived from up to four different strands have been found in the literature [180, 184, 204, 270].

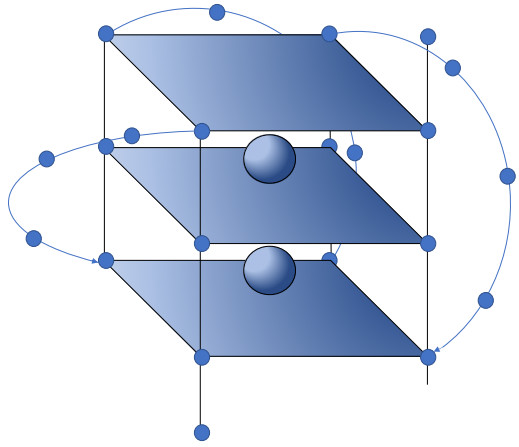

Figure 2: G-quadruplex structure formed by three stacked G-tetrads, metal cations, and side loops.

1.2 RELATION BETWEEN CANCER AND G-QUADRUPLEXES

GQ's have been associated with a large number of genomic functions, such as transcription and replication [111], and studies have revealed their presence in key regions of the human genome, such as promoters [228], untranslated regions (UTRs) [29], and telomeres [34]. In the case of telomeres, the formation of telomeric GQ's has been shown to inhibit telomerase activity [71], which is responsible for maintaining the length of telomeres and it is involved in 85% of cancers [161]. This inhibition causes apoptosis of cancer cells and, since telomerase is overexpressed in the majority of cancer cells and in relatively few somatic cells, it is recognized as a potential cancer-specific target that would avoid the death of somatic cells in chemotherapy treatments for tumors. For this reason, this strategy has been considered in current drug design to devise alternative therapies for cancer treatments, and

the studies on small molecules that bind and stabilize GQ's have become a current hot topic of pharmaceutical and medical research [70].

1.3 STABILIZATION OF THE G-QUADRUPLEX DNA STRUCTURE

To both carry out the inhibition of telomerase and alter the expression of oncogenes via a mechanism similar to that described above, GQ's have to be stabilized to ensure the disruption of these processes. Many organic small planar molecules [6] and some metal complexes [48] showed not only high affinity for GQ's, but also more selectivity for binding to GQ's than to dDNA, indicating that they may further stabilize the GQ's. They may interact with GQ's by means of end-stacking, groove binding, or the loop binding mode (Figure 3). The case of organic planar small molecules has been widely studied, and it has been shown that the following characteristics of the molecule increase the affinity for GQ's: (*a*) a π-delocalized system, which helps with the stacking of the G-tetrad; (*b*) a positive or partial positive charge, which interacts with the O6 of the guanines oriented toward the center of the G-tetrads; (*c*) positively charged substituents, which are able to interact with the negatively charged phosphate backbone; and/or (*d*) a surface area similar to the G-tetrads, which creates a greater affinity toward the GQ than toward regular DNA. Taking these characteristics into account, most of the studied ligands are organic molecules with a high number of aromatic rings, such as porphyrins, acridines, acridones, quinacridines, anthraquinones, porphyrazines, perylenes, quinoanthroxazines, barberines, bistriazoles, or coronenes. However, metal complexes offer a broad range of structural and electronic properties that can be exploited when designing new drugs. The metal center may be used as a structural focus to organize ligands in specific geometries and orientations for optimal GQ binding. In addition to their structural characteristics, the electronic properties of metal atoms may reduce the electron density on coordinated aromatic ligands. This

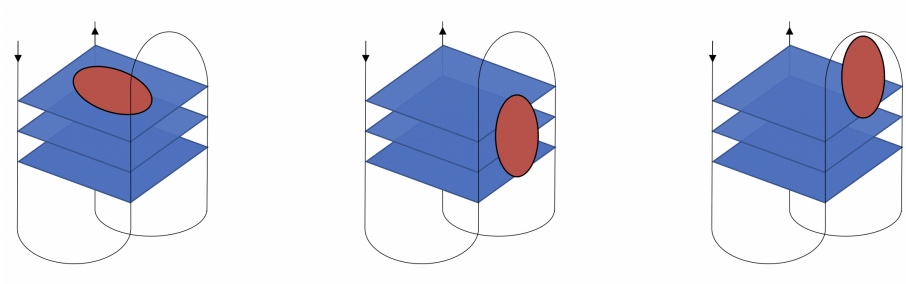

Figure 3: Representation of different binding modes between a ligand and a G-quadruplex DNA. (a) End stacking. (b) Groove binding mode. (c) Loop binding mode

allows the introduction of electron-poor systems, which are expected to show stronger π interactions with G-tetrads. Thus, although most of the reported molecules are purely organic structures, the interest in metal complex systems has increased during recent years.

1.4 COMPUTATIONAL METHODS AND STATE OF THE ART

Different molecules, binding modes, and mechanisms of stabilization of GQ's have been studied during the past decades with the aim of understanding and rationalizing the formation and stabilization of these GQ's. In many of these investigations, experimental studies have been supported by computational methods for a better understanding of the formation of GQ's or their interaction with ligands. Theoretical and computational methods used to address these systems can be divided into two types: (*a*) non-quantum methods and (*b*) quantum methods. The former allows long-timescale Molecular Mechanics (MM) simulations, enhanced simulation techniques such as metadynamics, and the use of coarse-grained force fields. The latter can be of different accuracy and efficiency and includes Semi-Empirical (SE), Quantum Mechanics/Molecular Mechanics (QM/MM), and Density Functional Theory (DFT) in reduced models. These types of methods can explicitly

treat the electronic structure of the system. However, they have a limited application in terms of dynamic information, since, in general, they are computationally demanding and only allow short-time simulations. Therefore, they are used to provide a more static picture of the different interactions in the systems, characterized by accurate electronic structure determination.

The most popular option for studying systems and processes involving DNA GQ's includes the use of Molecular Dynamics (MD) simulations, for two main reasons. First, most ligand–DNA interactions are non-covalent interactions and proceed without the break and formation of covalent bonds. Therefore, the use of MD simulations is, in principle, adequate. Second, GQ systems are relatively big, as are most biological systems. The use of MD simulations for this kind of systems is common in the scientific community because, in exchange for losing precision, it allows one to study the entire system, taking into account evolution over time. When carrying out MD simulations, the Amber and GROMACS software with parm99 and parmbsc0 force fields have generally been used. However, we must be careful when using force fields, because each one is parameterized to describe a type of system. In general, the force field that has given the best results has been the parmbsc0. It must be said that these MD simulations may be complemented with QM/MM calculations to give a better description of the electronic structure of some important parts of the system and the interactions and/or processes occurring during the stabilization of GQ's. The QM/MM methods are considered to be the state of the art for systems in which a ligand interacts with DNA. MD simulations have been used in several studies aiming at reproducing the conformational variety of the loops. More work is needed in this area because of the lack of parameters and the performance of the current force fields. Šponer and colleagues [80, 98, 112, 114, 195, 230, 272] have developed very useful and elegant work during the past years with the development and reparameterization of force fields, especially in the case of torsion

parameters. MD simulations have also been the first choice to study the processes of folding and unfolding of these non canonical secondary structures of DNA, in which the proper description of weak interactions plays a crucial role. As for the study of the loops, the choice of the force field is critical for a good description of these processes, and again, long time simulations are crucial for a proper description of the folding and unfolding processes in GQ's; the use of not only GPUs but also alternative methods like metadynamics and coarse-grained methods accelerate the classical and standard MD simulations. MD simulations have been used to study the stabilization of GQ's through interactions with ions and small molecules such as organic ligands and metal complexes. During the last years, many works were published aiming at the comprehension and rationalization of the stabilization of GQ's through the interaction with ions and small molecules. These kinds of studies, involving organic ligands and metal complexes, are very interesting from the pharmaceutical and medical points of view because, using the results and conclusions of these works, we could devise new small molecules employing substitution of ligands and changes in metals to improve the interactions and stabilization of GQ's. Thus, studies in the future on derivatives of these organic ligands and metal complexes could be very useful for pharmaceutical and medical applications. In this sense, an important aspect to take into account is not only the affinity of these small molecules with GQ's, but also the selectivity favoring their interaction with GQ's versus dDNA. Affinities and selectivities of small molecules with dDNA and GQ's have been analyzed with MD simulations and QM/MM approaches; noteworthy examples include the computational works of Barone and colleagues [36, 37, 85, 240, 241] and the experimental works of Neidle and colleagues [6, 56, 90–92, 113, 180], in which synergies with computational studies were found.

The state of the art on computational studies of GQ's have not been limited to the MD and QM/MM simulations; Density Functional The-

ory with dispersion corrections (DFT-D) calculations on GQ's have also been conducted on reduced models. These DFT-D methods can describe explicitly the electronic structure to give a more detailed picture of the molecule. However, they cannot deal with many atoms and the evolution of the system over time, which makes them incapable of describing the system in its complexity. This DFT-D choice was used especially by Fonseca Guerra and colleagues [88, 89, 260, 266, 267], who used reduced models of different numbers of G-tetrads to study the role of the ions in the channel, the stability of different kinds of GQ's (guanine tetrads versus adenine tetrads), and the cooperativity of the weak non-covalent interactions (hydrogen bonds and stacking). One of the main conclusions obtained with this DFT-D approach was that alkali cations are not mandatory for the stabilization of GQ's. However, their interactions with the DNA bases of the tetrads, when such alkali cations are inside the ion channel, give some extra stability to the system that helps to maintain the non-canonical secondary DNA structure. Another interesting finding obtained with reduced models and DFT-D computations was that RNA-GQ's may be more stable than the DNA-GQ's, a result that seems to be related to an extra hydrogen bond involving the 2-OH of the ribose of the RNA with the phosphate oxygen atoms, which gives higher conformational stability to the structure. Finally, it is still challenging the use of SE methods to large biological systems, even though once dispersion corrections on the methods have been introduced. Unfortunately, few works have tried to apply SE methods to the study of GQ's. These studies only calculated charges to be applied with other methods or studied the interaction of GQ's with some small molecules. However, some work by Dinçalp et al. used SE methods to study spectroscopic properties [60].

1.5 LINEAR-SCALING COMPUTATIONAL METHODS

The increasing computational power and the developments on computational techniques allows to study larger molecular systems. In this sense, properly parallelized linear-scaling methods allow the study of large molecular systems in a reasonable computational time. These methods allow to study explicitly the electronic configuration for the entire system. There are few works that describe large biological systems using linear-scaling DFT methods, such as the works done for the DNA [178, 179, 227]. On the other hand, a near linear-scaling Coupled Cluster method has been developed recently, which is called Domain-based Local Pair Natural Orbital (DLPNO)-Coupled-Cluster with iterative inclusion of single and double excitations with perturbative inclusion of triple excitations (CCSD(T)). This method is a very efficient method, which allows the application of the Coupled Cluster methods to large biological systems [199, 200].

1.6 OBJECTIVE OF THIS THESIS

GQ's have become a topic of great interest during the last decade. Studying and understanding the biological mechanisms that govern the formation and stabilization of these secondary DNA structures and their interaction with small molecules has become of vital importance since they offer an alternative to current, too toxic, cancer treatments. This thesis is a contribution to the field of study of GQ's through computational techniques. The thesis address the following issues:

- The use and analysis of the accuracy and performance of semi-empirical, Linear-Scaling DFT, and near Linear-Scaling DLPNO-CCSD(T) methods, for the study of G-quadruplexes and their interaction with small molecules.

- To study and rationalize the nature of the interaction between the G-quadruplex and the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex, which has been synthesized and shown to have cytotoxic effects against several tumoral cell lines.

- To perform a study to rationalize the binding affinity and selectivity that the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex has towards the dDNA and G-quadruplex secondary DNA structures.

2

METHODS

2.1 QUANTUM THEORY

Computational chemistry and molecular modeling are used to characterize and to predict the behavior of molecules. The most precise method to calculate the energy of molecular systems is by solving the Schrödinger equation, the basis of the Quantum Mechanics (QM). This theory, developed in the twenties of the last century [61, 101, 215, 216], changed the viewpoint of physicist over the microscopic world. Bohr explained the stability of the simplest atom, hydrogen, by fixing stationary orbits of the electrons around the nucleus. At the beginning, the theory was mainly the playground of physicists, but it soon found applications in chemistry, giving rise to quantum chemistry. In QM the electrons in a molecule are treated explicitly, which is computationally expensive. Due to this, the method has been traditionally limited to systems with a small number of atoms. Over the years quantum chemistry has produced tools and developed methods in order to calculate, understand and predict molecular properties trying to reduce the computational cost while keeping the accuracy.

For large molecules as DNA, the description of the system requires simpler methods such as classical molecular mechanics (MM). These methods are based on classical principles of motion, evaluating the energy as the function of the position of the atoms. The electrons are not treated explicitly in this type of methods.

In QM, in principle, all the information of a system, a molecule, for example, may be obtained from the *wavefunction*, Ψ. In order to deter-

mine it, we need to solve the Schrödinger equation (here in its time independent form):

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

where the Hamiltonian, $\hat{H}$, contains the kinetic energy terms of the electrons and the nuclei, the interactions between the nuclei, the electrons, and between nuclei and electrons. When applying $\hat{H}$ to the wavefunction, Ψ, the energy of the system E is obtained, times the wavefunction itself. For a system with N_e electrons and N_N nuclei, the Hamiltonian in atomic units would be expressed as in Equation 2.2, aside from relativistic effects[1] and spin-orbit coupling[2].

$$\begin{aligned} \hat{H} &= \hat{T}_e + \hat{T}_N + \hat{V}_{e-e} + \hat{V}_{e-N} + \hat{V}_{N-N} \\ &= -\sum_i^{N_e} \frac{\nabla_i^2}{2} - \sum_A^{N_N} \frac{\nabla_i^2}{2M_A} \\ &\quad -\sum_A^{N_N}\sum_i^{N_e} \frac{Z_A}{r_{iA}} + \sum_{i>j}^{N_e} \frac{1}{r_{ij}} + \sum_{A>B}^{N_N} \frac{Z_A Z_B}{R_{AB}} \end{aligned} \tag{2.2}$$

In Eq 2.2 $\hat{T}_e$ and $\hat{T}_N$ are the operator related to the kinetic energy of electrons and nucleus; $\hat{V}_{e-e}, \hat{V}_{e-N}, \hat{V}_{N-N}$ terms represent, respectively, the electron-electron, electron-nucleus and nucleus-nucleus interactions. M_A is the mass of the nucleus A, Z_A and Z_B are atomic number of nucleus A and B, respectively; and $R_{AB} = \left|\vec{R}_B - \vec{R}_A\right|$, $r_{ij} = \left|\vec{r}_j - \vec{r}_i\right|$, $r_{iA} = \left|\vec{R}_A - \vec{r}_i\right|$ are the distances between nuclei A and B, electrons i and j and electron i and nucleus A, respectively.

Eq 2.2 can be simplified with the well known Born-Oppenheimer approximation[38], which states that since the nuclei are heavier than

1 When a mass moves at a speed approaching the speed of light, its inherent properties are changed in several ways; these are known as the Relativistic Effects which for the quantum description of the heaviest atoms in the periodic table may become crucial

2 Magnetic interaction between the spin of the electron and the angular momentum of the orbital

the electrons, their movement will be slower, hence the nuclear kinetic term, $\hat{T}_N$, can be neglected. As a consequence the nucleus-nucleus potential energy term, $\hat{V}_{N-N}$ can be considered as a constant, because it only depends on the inter-nuclear distances, which are fixed. The Born-Oppenheimer approximation results in the approximate Hamiltonian shown in Equation 2.3.

$$\begin{aligned} \hat{H}_e &= \hat{T}_e + \hat{V}_{e-e} + \hat{V}_{e-N} + \hat{V}_{N-N} \\ &= -\sum_i^{N_e} \frac{\nabla_i^2}{2} - \sum_A^{N_N} \sum_i^{N_e} \frac{Z_A}{r_{iA}} + \sum_{i>j}^{N_e} \frac{1}{r_{ij}} + \sum_{A>B}^{N_N} \frac{Z_A Z_B}{R_{AB}} \end{aligned} \tag{2.3}$$

Here $\hat{H}_e$ is the electronic Hamiltonian operator. Therefore, the electronic wavefunction, Ψ_e, depends explicitly on the electron coordinates, and parametrically on the nuclei position. It is worth mentioning that for any arrangement of nuclei, one obtains an electronic energy, and also a nucleus-nucleus interaction. This builds up an effective $3N_N$-dimensional energy potential (for N_N nuclei), the topology of which will provide the energies of different conformations, and the paths through which two energetically local minima structures are connected, which is often called Potential Energy Surface (PES). The PES of a set of atoms is a feature that provides the information needed to discern the mechanisms of chemical transformations, and their energies.

Many mathematical methods have been developed to provide a way to find the electronic wavefunction for any given nuclear arrangement. One of the first developed method was Hartree-Fock (HF) [72, 97], where the wavefunction is an antisymmetrized product of one-electron orbitals. The electrons are treated as moving in a mean field due to the nucleus and the other electrons. There are different ways in which the correlation can be taken into account. One of them corresponds to the perturbational methods such as n^{th} order Møller-Plesset theory (MPn hereafter) [164]. In this method the electron correlation is treated as

a perturbation of the HF problem. In the Configurational Interaction method (CI) [221, 222] the wavefunction is expressed as a linear combination of Slater determinants to provide a variational solution to the exact many-electron wavefunctions. There are other methods such as Coupled Cluster (CC) [24–26, 186], Complete-Active-Space Self Consistent Field (CASSCF) [3, 4] or the Multi-Reference Configuration Interaction (MRCI) [130, 258]. These methods are very useful and accurate tools to study the electronic properties of both ground or excited states.

There are other methods such as the semi-empirical (SE) methods, which are much faster but less accurate due to the fact that they are less rigorous in their treatment of the underlying physics involving many approximations and parameters from empirical data. A more accurate approach is DFT [105, 131, 181, 182] method, very closely related to HF (simplest *ab initio* method), but modified in a way that it takes into account the electronic correlation, hence its accuracy is highly improved.

2.2 SEMI-EMPIRICAL METHODS

The SE methods are a simplification of the HF method, to increase its performance. In SE methods only the valence electrons are explicitly considered, the core electrons are taken into account by introducing functions to model the combined repulsion caused by the nuclei and core electrons. Furthermore, only a minimum number of exponential *Slater type orbitals* (STO) functions are used as basis set in order to accommodate the valence electrons.

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_l^m(\theta,\varphi)r^{n-l}e^{-\zeta r} \tag{2.4}$$

where N is a normalization constant, Y_l^m are the spherical harmonic functions, n, m and l are quantum numbers, θ is the polar angle and φ

is the azimuthal angle, r is the distance of the electron from the nucleus and ζ is the constant related to the effective charge of the nucleus.

Many of the SE methods uses the Zero-Differential Overlap approximation (ZDO) which simplifies the problem by ignoring electron repulsion integrals. All two-electron integrals involving three- or four-center are neglected, while the remaining integrals are made introduced as parameters, which are fitted to reproduce experimental data, or high level calculations. That is why these methods are called SE. The different SE methods differ on how many integrals are neglected and how the parametrization is done such as Neglect of Diatomic Differential Overlap (NDDO), Intermediate Neglect of Diatomic Differential Overlap (INDO) or Complete Neglect of Differential Overlap (CNDO). In this work, NDDO model and its different improved versions have been used.

2.2.1 *Neglect of Diatomic Differential Overlap Approximation (NDDO)*

In this aproximation, all the terms from the two atomic orbitals which are on different centers are set to zero, resulting in the following overlap integral (Eq 2.5) and one-electron operator (Eq 2.6) equations:

$$S_{\mu\upsilon} = \langle \mu_A | \upsilon_B \rangle = \delta_{\mu\upsilon}\delta_{AB} \tag{2.5}$$

$$\hat{h} = -\frac{1}{2}\nabla^2 - \sum_A^{N_N} \frac{Z'_A}{\left|\vec{R_A} - \vec{r}\right|} = \frac{1}{2}\nabla^2 - \sum_A^{N_N} V_A \tag{2.6}$$

Modified methods used molecular experimental data for fitting the parameters. MNDO, AM1 and PM methods are derived from the

NDDO approximation and differs in the way they treat the core-core repulsion and how the parameters are assigned.

2.2.2 *Parametric Method Models (PM)*

After the development of AM1, improvements were made to the method for parameter optimization. The result was the Parametric Method Number 3 (PM3), developed by Stewart [231–233]. In the PM3 method, the AM1 expression for the core-core repulsion was kept, but in this case two Gaussian core-core functions were assigned to each atom (Eq. 2.7).

$$E_N^{MNDO}(A,B) = E_N^{MNDO}(A,B) + \frac{Z'_A Z'_B}{R_{AB}} \times \left[\sum_k \alpha_{kA} e^{-b_{kA}(R_{AB}-c_{kA})^2} + \sum_k \alpha_{kB} e^{-b_{kB}(R_{AB}-c_{kB})^2}\right] \tag{2.7}$$

Moreover, the parametrization strategy was different. While the AM1 was parametrized mostly by using small number of atomic data, PM3 was parametrized to reproduce a large number of molecular properties. PM3 treats hydrogen bonds rather well, but it amplifies the hydrogen-hydrogen attractions in many cases, resulting in serious problems when analyzing intermolecular interactions or conformations of flexible molecules. In addition, with only *s*- and *p*-functions taken into account, the method showed limitations to treat a large part of the periodic table. It is known that *d*-orbitals significantly improve the results for components of the second row elements, specially the ones which have the potential to be hypervalent. The main problem of taking into account *d*-orbitals arises in the increase of two-electron integrals. However, Thiel and Voityuk showed that the large accuracy obtained when adding *d*-orbitals outweighs the extra computational cost [242–244].

After several modification of the NDDO core-core interaction term, the method of parameter optimization and the addition of *d*-orbitals resulted in the PM6 method. The general form of the core-core interaction used in PM6 is given in Eq. 2.8.

$$E_N^{PM6}(A,B) = Z'_A Z'_B \langle s_A s_A | s_B s_B \rangle \left[1 + \chi_{AB} e^{-\alpha_A (R_{AB} + 0.0003 R_{AB}^6)}\right] \quad (2.8)$$

The following improvements on SE methods targeted the non-covalent interactions, especially hydrogen bonds. Different corrections were made to the PM6 method, resulting in different versions such as PM6-D3 [84], PM6-DH [196], PM6-DH+ [133], PM6-DH2 [135], PM6-D3H4 [198], PM6-DH2X [197], and PM6-D3H4X [41]. The corrections can be split into two main groups: Dispersion, H-bond, and halogen bond corrections. These interactions are essential in order to reach the so-called chemical accuracy, and for biological applications.

The dispersion correction applied to the PM6 method was taken from the Grimme's D3 dispersion correction developed for DFT-D, which had been tested successfully on thousands of different systems including inter- and intramolecular cases.

On the other hand, there are three different Hydrogen-bonding correction generations.

The first-generation correction [135] made use of charges q on the acceptor (A) and hydrogen (H) atoms, the H-bond distance r and a cosine term that promotes a 180º angle situation for the A-H-D angle. The main problem of this correction was that it was only dependent on the H-bond distance, leading to discontinuous potentials around 90º for the A-H-D angle. These errors were fixed in the second-generation corrections.

The second-generation H-bonding correction took the following form [135]:

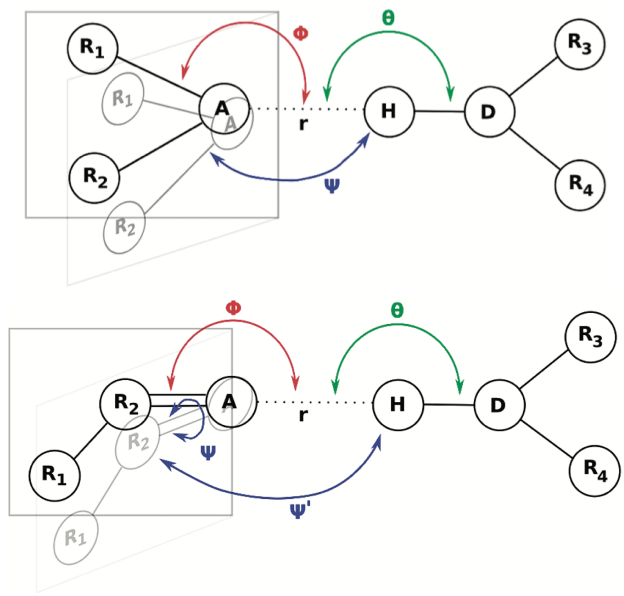

Figure 4: Illustration of the geometric features of hydrogen-bonding in the second-generation corrections.

$$E_{H-bond} = \left[a\frac{q_A q_H}{r^b} + cd^r\right]\cos(\theta)\cos(\phi)\cos(\psi) \tag{2.9}$$

The correction was included in the PM6-DH2 method, used in this work. It can be defined with six internal coordinates (see Figure 4) which defines better the H-bonding nature and leads to a more robust method.

The third-generation corrections [133] takes the H-bond as a charge-independent atom-atom term between two atoms capable of serving as an acceptor or donor (i.e., O, N), weighted by a function, f_{geom} (Eq. 2.10), that depend on the relative orientation of the two fragments, and which includes a dumping function to correct the short- and long-range behaviors:

$$f_{geom} = \cos(\theta_A)^2\cos(\phi_A)^2\cos(\psi_A)^2\cos(\phi_B)^2\cos(\psi_B)^2 f_{bond} \quad (2.10)$$

The last improvements on PM methods were compiled in the latest PM7 method [235]. It was reparemetrized using experimental and high level *ab initio* reference data. Moreover, two changes were made to the method: a modification to improve the description of non-covalent interactions, and two minor errors in the NDDO formalism were also corrected. Besides, the reduced or missing repulsion between certain atom pairs: Na-Na, Br-N, Br-O, Br-Br, S-N, S-S, S-O, S-Cl, I-N, I-O, and I-I. In addition, some changes were applied [234], adding the consideration of a constrain on the value of the core-core interaction and the electron-electron repulsion integral to allow some terms to converge to exact values at separations greater than 7 Å, which allows a better treatment of solid state.

2.3 DENSITY FUNCTIONAL THEORY METHODS

The Density Functional Theory formalism replaces the N-electron wave function and associated Schrödinger equation by the much simpler electron density $\rho(\overrightarrow{r})$ which is a function of the three spatial variables. Then the electronic state, the energy and all the electronic properties of any system can be described in terms of this $\rho(\overrightarrow{r})$ [66, 182].

Hohenberg and Kohn [105] proved that the electronic properties of a system with a non-degenerate ground state are uniquely determined by the electron density. Hence, the ground-state energy E_o is a function of $\rho(\overrightarrow{r})$. They also established an energy variational principle for the energy functional, analogous to the variational principle for wave functions. Thus, knowing the exact form from the $E[\rho]$ functional, we can search for the ground state density (as it is the case of the wave function). In the Kohn-Sham version of DFT they showed [131] that

the electronic energy E_o of a N-electron molecule with a ground-state electron probability density ρ can be written by

$$\begin{aligned} E_0 &= -\tfrac{1}{2}\sum_{i=1}^{N}\left\langle \psi_i(1) \mid \nabla_i^2 \mid \psi_i(1)\right\rangle + \int v(r)\rho(1)d\vec{r_1} \\ &\quad +\tfrac{1}{2}\int\int \frac{\rho(1)\rho(2)}{r_{12}} d\vec{r_1}\, d\vec{r_2} + E_{xc}[\rho] \end{aligned} \tag{2.11}$$

where $v(r) = -\sum_\alpha \frac{Z_\alpha}{r_{1\alpha}}$ is the external potential due to the nuclei, ψ_i are the Khon-Sham orbitals, and the $E_{xc}[\rho]$ is the exchange-correlation energy.

Density functional methods have proved to give excellent results in most chemical systems [140], with results comparable to those given by CPU intensive electron-correlation methods. However, they frequently overestimate bond dissociation energies [250]. The hybrids of HF and DFT theories increment the accuracy of the dissociative energies as was validated by Johnson *et al.* [116].

One of the main advantage of these methods is that with a similar computational cost as for HF methods, they include some kind of electron correlation, being the major drawback that the correlation effects cannot be sorted out precisely. They are already mixed from the beginning with the uncorrelated solution. Despite of these facts, DFT has been found to yield good results for ground state properties of various chemical systems, with a quality comparable to MP2 results [271] or even better in some cases. Due to their relative low computational cost, DFT is the method of choice for large systems, for which the inclusion of electron correlation by MP or CI methods is prohibitive. However, one should always be careful to benchmark the specific functional, used in a given application.

2.3.1 *Linear-Scaling Density Functional Theory*

The application of traditional DFT to problems in fields such as biology is hindered by the scaling of the computational effort with the system size. In many cases, for such kind of systems one needs an accurate description of structures with more than several hundreds or thousands of atoms. This calculations are too expensive using standard DFT calculations, since they scale as $\mathcal{O}(N^3)$, where N is the number of atoms. To overcome this problem, methods that scale linearly $\mathcal{O}(N)$ (linear-scaling), or large-scale techniques have been developed [81]. Moreover, the growth in computer power comimg from the increase in the number of processors or cores, makes mandatory a good parallel efficiency. In general terms, linear-scaling methods need to overcome two key points to perform efficient DFT calculations on large-scale systems and massively parallel computers:

- To develop a method to calculate the electronic structure suitable for massively parallel calculations.
- To solve the electronic problem with better scaling than $\mathcal{O}(N^3)$, ideally with linear scaling $\mathcal{O}(N)$.

There are several ways to achieve these goals such as the use of the iterative diagonalization technique [138, 185], fast Fourier transforms (FFTs) [55, 173], the use of *ab initio* psuedopotentials to replace the core electrons [10, 192, 249, 254], or in the case of plane-waves, the reduction of the number of plane-waves by using augmented wavefunctions in core regions (as in the linearized augmented plane-wave LAPW) [35, 156, 225, 226], or the projector augmented wave (PAW) [104, 139] methods. The main codes that include fully or partial linear-scaling and large-scale functionality are, in alphabetical order: CONQUEST [46], ErgoSCF [63], FEMTECK [251, 252], FreeON [67], ONETEP [174], OpenMX [175], PROFESS [110], and SIESTA [207].

2.3.2 *The SIESTA Method*

In this work, we have manly used the SIESTA code. The SIESTA project (Spanish Initiative for Electronic Simulations with Thousands of Atoms) [7, 176, 177, 213, 229] started in 1995 and it is one of the earliest linear-scaling codes widely available, which makes use of a fully self-consistent DFT, based on a linear combination of atomic orbital (LCAO) basis sets. Apart from Born-Oppenheimer, the most basic approximations concern the treatment of exchange and correlation (XC), the use of standard norm-conserving pseudopotentials [11, 93] in their fully non local form [129], and the use of numeric atomic orbitals, which are strictly localized and vanishes beyond a given r_c cutoff radius. In principle, this fact makes possible in principle an $\mathcal{O}(N)$ scaling, which lead to a general-purpose flexible LS-DFT method. However, it must be mentioned that even though the matrix elements are computed with linear-scaling algorithms, the diagonalization is porportional to $\mathcal{O}(N^3)$.

2.3.2.1 *Pseudopotential*

When solving the electronic structure of any system, all electrons must be included since they contribute to the potential experienced by other particles. In practice, the core electrons of an atom are weakly perturbed by chemical changes in comparison to the valence electrons, and therefore, several approximate methods have evolved alternatively to treat these core states to reduce computational expense. The core electrons and the protons in the nucleus, which have opposite charges and therefore partially cancel each other, can be replaced by a combined effective potential, known as a pseudopotential. The pseudopotential replaces the exact potential due to the nucleus and core electrons, within a given radius of the atomic center, by an effective potential. Within this distance, known as the *core radius*, the potential is

smoothed and tends to a finite value at the nucleus while matching the true potential at the boundary. Due to the smoothing of the potential, the radial nodes of the valence states are lost in the core region since there is no longer any requirement to maintain orthogonality to the core states. In most of the cases a non-local pseudopotential is used, which implies that there is a different potential for each angular momentum channel, with a separate core radius, r^{l}_{core}, appropriate to that channel.

SIESTA reads the first-principles norm-conserving pseudopotentials in semi-local form [93], with a different radial potential $V_l(r)$ for each angular momentum l, generated scalar-relativistically [11, 128]. Generally, the Troullier-Martins parametrization is used to smooth the pseudopotential, reducing the computational resources needed [249]. This semi-local form is then transformed into the fully non-local form proposed by Kleinman and Bylander (KB) [129].

2.3.2.2 *Basis set*

Numerical solution of the Kohn-Sham equations is performed by expanding the orbitals in terms of a computationally convenient mathematical function: the basis set. The coefficients that determine how much these functions contribute are found by applying the variational principle. In choosing the optimal basis set for large linear-scaling calculations, we are guided by the need for locality in real space and the requirement to minimize the number of basis functions needed to obtain reasonable numerical precision. If the pseudopotentials of the form described in the previous section are employed, then neither existing Slater or Gaussian basis sets will be of the correct form, due to the modification of shape in the nuclear region.

In the SIESTA methodology, the standard choice of basis set is pseudoatomic orbitals (PAOs), φ^{A}_{nml} for atom I, which are tabulated on a

logarithmic radial grid for each angular momentum and then multiplied by appropriate spherical harmonics:

$$\varphi^{I}_{lmn}(r,\theta,\varphi) = R^{I}_{nl}(r)Y_{lm}(\theta,\varphi) \tag{2.12}$$

While the PAOs above decay rapidly with distance, as do other atomic-centered basis functions, they only tend asymptotically to zero at infinite radius. To achieve linear-scaling it is necessary to impose on the Hamiltonian strict locality in real space. Following the Sankey and Niklewski method [214], the eigenfunctions of the pseudoatomic problem are found within the confines of a spherical boundary at which the potential becomes infinite. In this way, the tails of the PAOs are modified such that they go rigorously to zero at a given radius. This radius, r_c, can be selected to be different for each angular momentum.

2.3.2.3 *The LMKLL van der Waals density functional*

The van der Waals (vdW) interaction is a quantum mechanical phenomena, and the understanding and description of such force is critical for biostructures (e.g. DNA and protein structures). The vdW force at one point thus depends on charge events at another region and is a truly non-local correlation effect. In this thesis the vdW LMKLL functional has been used, which treat dispersion implicitly, unlike the usual Grimme dispersion corrections [82, 84], which are geometry-dependent. The key to the vdW-DF method is the inclusion of a long-range piece of the correlation energy $E_c^{nl}[\rho]$, a fully non-local functional of the density ρ (eq. 2.13).

$$E_c^{nl}[\rho] = \int d^3r \int d^3r' \rho(r)\phi(r,r')\rho(r') \tag{2.13}$$

2.4 DOMAIN-BASED LOCAL PAIR NATURAL ORBITAL COUPLED CLUSTER WITH SINGLE, DOUBLE, AND PERTURBATIVE TRIPLE EXCITATIONS (DLPNO-CCSD(T))

The coupled cluster (CC) method is an approximation to solve the time independent Schrödinger equation for electrons. It is widely used since its introduction into quantum chemistry by Čížek and Paldus [54], for high-accuracy calculations.

Coupled-cluster theory provides the exact solution to the time-independent Schrödinger equation. The wavefunction of the coupled-cluster theory is written as an exponential ansatz:

$$|\Psi\rangle = e^{T} |\Phi_0\rangle = \Omega |\Phi_0\rangle \tag{2.14}$$

where $|\Phi_0\rangle$ is the reference wavefunction, which is typically a Slater-type determinant constructed from Hartree-Fock molecular orbitals (MOs). The cluster operator T is composed of a series of connected operators that can be expanded in terms of its components that introduce single Φ_i^a, double Φ_{ij}^{ab}, triple Φ_{ijk}^{abc}, etc. excitations into the wavefunction.

The Schrödinger equation can be written using coupled cluster wave functions as,

$$H |\Psi_0\rangle = H e^{T} |\Phi_0\rangle = E e^{T} |\Phi_0\rangle \tag{2.15}$$

Using a "projective" technique, one may left-multiply this equation by the reference, Φ_0, to obtain an expression for the energy,

$$\left\langle \Phi_0 \left| \hat{H} e^{T} \right| \Phi_0 \right\rangle = E \left\langle \Phi_0 \left| e^{T} \right| \Phi_0 \right\rangle = E \tag{2.16}$$

One of the most used coupled cluster methods is the CCSD(T), where there is a full treatment of singles and doubles excitations, whereas triples contribution is calculated non-iteratively using many-body perturbation theory.

2.4.1 *Domain Local Pair Natural Orbitals (DLPNO)*

The only obstacle when using CC methods in computational chemistry is the high computational cost of the calculation. The disk space requirements scale as a fourth power of the molecular size, and computer CPU requirements for Coupled-Cluster with iterative inclusion of single and double excitations (CCSD) and CCSD(T) calculations are well known to scale as $O(N^6)$ and $O(N^7)$, respectively. Over the years many approximations to the CC equations have been suggested and in one way or another most of them make use of localization techniques [8, 9, 50, 206, 238]. Among all the approximations, there is one variant developed by Pulay and co-workers, based on the concept of local correlation domains [39, 190, 208, 209] that has been further developed and implemented by Werner and co-workers [95, 193, 217]. In these methods, one uses localized internal orbitals and projected atomic orbitals (PAOs) to span the virtual space. With proper programming and extensive threshold, this method can be devised to be linear-scaling with respect to all computational resources (disk space, main memory, and CPU). With the standard definition of the correlation domains introduced by Boughton and Pulay [39], the local correlation approaches recover up to 99.6% of the correlation energy on average [259].

In the DLPNO-CCSD(T) method the local correlation problem follows a different strategy. While the internal space is also comprised of localized MOs, it spans the external correlating space by pair natural orbitals (PNOs) [168]. The PNOs are a highly compact set of orbitals that are different for each electron pair. In the PNO method, one constructs

approximate natural orbitals for each electron pair in order to obtain rapidly converging expansion of the dynamic correlation energy.

2.5 ENERGY DECOMPOSITION ANALYSIS (EDA)

The Energy Decomposition Analysis (EDA) focuses on the intrinsic interactions A–B between the molecules A and B which must be distinguished from the bond dissociation energy (BDE). The latter term involves the geometrical and electronic relaxation of the fragments A and B from the molecule A–B in the equilibrium geometry of the electronic ground state. Thus, the EDA gives a faithful representation of the actual bonding situation in the molecule. Since the sum of the various energy terms of the EDA and the relaxation energy give the BDE, they are connected to an observable. This makes the EDA a very attractive and powerful tool for the interpretation and explanation of chemical phenomena such as molecular structure and reactivity. In principle, the EDA can be used in conjunction with any quantum chemical approach. Herein, it is combined with the quantitative MO model contained in the Kohn–Sham approach to DFT, which includes correlation energy. For *ab initio* methods, the EDA has only been developed for Hartree–Fock calculations until today. This is the reason that the EDA method is usually employed in conjunction with DFT calculations, although there are also other ways to carry out EDAs, such as the one developed by ORCA through the DLPNO-CCSD.

The focus of the bonding analysis in the EDA approach is the instantaneous interaction energy ΔE_{int} of a bond A–B between two fragments A and B in the particular electronic reference state and in the frozen geometry of AB. The interaction energy is divided into three main components:

$$\Delta E_{int} = \Delta E_{elstat} + \Delta E_{Pauli} + \Delta E_{orb} \tag{2.17}$$

The term ΔE_{elstat} corresponds to the electrostatic interaction between the unperturbed charge distributions of the atoms, which it is usually attractive. The Pauli repulsion ΔE_{Pauli} arises as the energy change associated with the transformation from the superposition of the unperturbed electron densities $\rho E(\alpha) + \rho E(\alpha)$ of the isolated atoms to the wavefunction. The orbital interaction ΔE_{orb} accounts for charge transfer (i.e., donor–acceptor interactions between occupied orbitals on one moiety with unoccupied orbitals of the other, including the HOMO-LUMO interactions), polarization (empty/occupied orbital mixing on one fragment due to the presence of another fragment), and electron-pair bonding (the stabilization arising from the formation of the electron-pair bonding).

There is one further type of interatomic interactions, dispersion interaction (ΔE_{disp}), which comes from the attractive forces between the induced dipoles of interacting species. Dispersion interaction between one pair is weak, which is the reason that it was often considered to be negligible for chemical bonding. However, if the interacting species are very large, the number of pair interactions grows rapidly and the associated attractive forces may provide an important part of chemical bonding. If an explicit correction term for dispersion interaction is employed such as in the methods suggested by Grimme [82, 83], the EDA results remain unchanged and the dispersion correction appears as an extra term ΔE_{disp}. If the dispersion interaction is part of the functional it will change the EDA results.

2.6 QUANTUM THEORY OF ATOMS IN MOLECULES (QTAIM)

The observation that some properties attributed to atoms and functional groups are transferable from one molecule to another has played a key role in the development of chemistry. This observation provides a basis for group additivity schemes and is exemplified by the constancy of group contributions to thermodynamic and spectroscopic

properties. Quantum Theory of Atoms In Molecules (QTAIM) [15], developed by Professor Richard F. W. Bader and his coworkers, relies on quantum observables such as the electron density, ρ, and energy of densities to address a partition on atomic basins of various molecule properties.

2.6.1 *The Topology of the Electron Density*

The topology of the electron density is dominated by the attractive forces of the nuclei leading to its principal topological feature, a substantial local maximum at the position of each nucleus.

A Critical Point (CP) in the electron density is a point in space at which the first derivatives of the density vanish, i.e.:

$$\nabla\rho = i\frac{d\rho}{dx} + j\frac{d\rho}{dy} + k\frac{d\rho}{dz} \rightarrow \begin{cases} = \overrightarrow{0} & (\text{At critical points} \\ & \text{and at } \infty) \\ \text{Generally} \neq \overrightarrow{0} & (\text{At all other points}) \end{cases} \tag{2.18}$$

where the zero vector signifies that each individual derivative in the gradient operator is zero. The gradient of a scalar function such as $\rho(r)$ (Eq. 2.18) at a point in space is a vector pointing at the direction in which $\rho(r)$ undergoes the greatest rate of change and having a magnitude equal to the rate of change in that direction. The maximum at the position of a nucleus constitutes one type of CP, namely, a Nuclear Critical Point (NCP).

2.6.2 *Critical point classification*

CP are classified according to their *rank* (ω) and *signature* (σ) and are symbolized by (ω,σ). The rank is the number of non-zero curvatures of ρ at the CP. A CP that has $\omega < 3$ is mathematically unstable and will vanish under small perturbation of the density caused by nuclear motion. The presence of such a CP (with a rank less than three) indicates a change in the topology of the density and, hence, a change in the molecular structure. For this reason, a CP with $\omega < 3$ is generally not found in equilibrium charge distribution and one nearly always founds $\omega = 3$. The signature is the algebraic sum of the signs of curvatures, i.e. each of the three curvatures contributes ± 1 depending on whether it is a positive or negative curvature. There are four types of stable CP having three non-zero eigenvalues:

- (3,-3) Three negative curvatures: ρ is a local maximum.
- (3,-1) Two negative curvatures: ρ is maximum in the plane defined by the corresponding eigenvectors but is a minimum along the third axis which is perpendicular to this plane.
- (3,1) Two positive curvatures: ρ is a minimum in the plane defined by the corresponding eigenvectors and a maximum along the third axis which is perpendicular to this plane.
- (3,3) Three curvatures are positive: ρ is a local minimum.

Each type of CP described above is identified with an element of chemical structure: (3,-3) Nuclear Critical Point NCP; (3,-1) Bond Critical Point (BCP); (3,1) Ring Critical Point (RCP); and (3,3) Cage Critical Point (CCP). Figure 5 shows the molecular graph (the set of bond paths and CP) of a cubane molecule. The bond critical path is a single line of maximum electron density linking the nuclei of two chemically-bonded atoms.

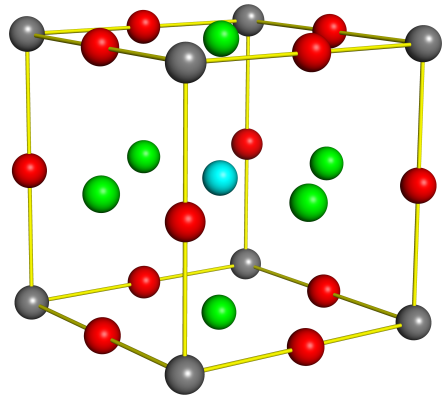

Figure 5: The molecular graph of a cubane molecule, showing the bond paths (lines) and the different CP: nuclear (Carbon atoms represented by grey dots), bond (red dots), ring (green dots), and cage (blue dot) CP.

2.6.3 *Bond properties*

A zero-flux surface is defined by a particular set of $\nabla\rho(r)$ trajectories all the members of which terminate at a single point, the BCP, where $\nabla\rho(r) = 0$. There is one BCP between each pair of atoms that are bounded. In addition to the set of trajectories which terminates at BCP and define an interatomic surface, a pair of trajectories originates at the BCP with each member of the pair terminating at one of the nuclei of the chemically bonded atoms. This latter pair of trajectories defines the bond path [13]. Chemical bonding interactions are characterized and classified according to the properties of the electron and energy densities at the BCP, collectively known as "bond properties":

- Electron Density at the BCP (ρ_b). The strength of a chemical bond, its bond order (BO), is reflected in the electron density at the BCP (ρ_b):

$$BO = [A(\rho_b - B)] \tag{2.19}$$

where A and B are constants which depend on the nature of the bonded atoms.

- Bonded Radius of an Atom (r_b) and the Bond Path Length. The distance of a BCP from nucleus A determines the "bonded radius" of atom A relative to the interaction defined by the BCP, and it is denoted $r_{b(A)}$. If the bond path is coincident with the internuclear axis, then the sum of the two associated bond radii, termed the bond path length, equals the bond length.

- Laplacian of the Electron Density at the BCP ($\nabla^2\rho_b$). The Laplacian at the BCP is the sum of the three curvatures of the density at the CP, the two perpendicular to the bond path, $\lambda 1$ and $\lambda 2$, being negative (by convention, $|\lambda 1| > |\lambda 2|$) whereas the third, $\lambda 3$, lying along the bond path, is positive. The negative curvatures measure the extent to which the density is concentrated along the bond path and the positive curvature measures the extent to which it is depleted in the region of the interatomic surface and concentrated in the individual atomic basins.

2.7 NON-COVALENT INTERACTION PLOTS (NCI)

Historically, Non-Covalent Interactions (NCI) have been decomposed into several distinct contributions: electrostatics, polarization, exchange, and charge-transfer effects [126]. A possible approach is studying NCI through the electron density $\rho(r)$ and its derived scalar and vector fields. Such approach shares the key concept of DFT: the electron density can be directly related to the energy, and therefore, we should be able to study interactions purely from the density as well.

The NCI analysis provides an index, based on the electron density and its derivatives, that makes possible the identification of non-covalent interactions [57, 118]. The NCI index is based on a 2D plot of the reduced density gradient, s, and the electron density, ρ, where

$$s = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla\rho|}{\rho^{4/3}} \tag{2.20}$$

When a weak interaction is present, there is a change in the reduced gradient between the interacting atoms, appearing the density between interacting fragments. Channels appear in $s(\rho)$ associated with each interaction. This fact is highlighted in Figure 6. When we search for the points in the real space giving rise to this feature, the non-covalent region clearly appears in the molecular complex (green isosurface in Figure 6d).

Deeper analysis of the electron density in the channels is required to determine the origin of these channels ($\pi - \pi$ intetractions, hydrogen bonds, etc.). The strength of the interaction is given by the electron density values within the channels. Both attractive and repulsive interactions (i.e., hydrogen-bonding and steric repulsion) appear in the same region of density/reduced gradient space. To distinguish between attractive and repulsive interactions, the second derivatives of the density along the main axis of variation must be computed.

Taking as reference the divergence theorem [5], the sign of the Laplacian ($sign\nabla^2\rho$) of the density indicates whether the net gradient flux of density is entering ($\nabla^2\rho < 0$) or leaving ($\nabla^2\rho > 0$) an infinitesimal volume around a reference point.

On the basis of the divergence theorem [5], the sign of the Laplacian ($sign\nabla^2\rho$) of the density determines whether the density is concentrated or depleted at that point, relative to the surroundings. As the sign is dominated by negative contributions from the nuclei [14], to distinguish between different types of weak interactions, contributions

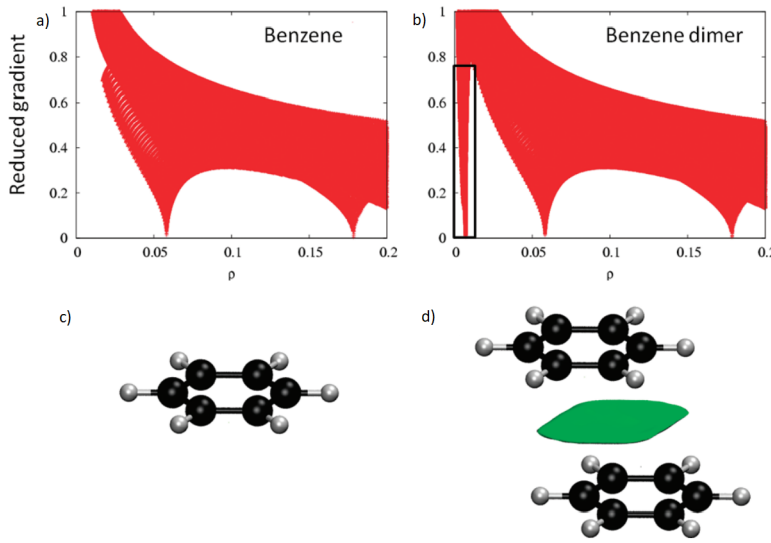

Figure 6: Comparison of the reduced density behavior for the benzene a) monomer and b) dimer; a singularity in s appears at low density values in the dimer case. c) Benzene monomer. d) Appearance of an intermolecular interaction surface in the benzene dimer, associated with the additional singularity in the $s(\rho)$) plot. The isosurface was generated for $s = 0.7$ au and $\rho < 0.01$ au. [57]

to the Laplacian along the axes must be analyzed. These contributions are the eigenvalues λ_i of the electron-density Hessian (second derivative) matrix, such that $\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$, ($\lambda_1 < \lambda_2 < \lambda_3$). The second eigenvalue (λ_2) can be either positive or negative, depending on the interaction type. Thus, analysis of the sign of λ_2 enables us to distinguish different types of weak interactions, while the density itself enables us to assess the interaction strength.

It is also possible to qualitatively classify interactions to complete the NCI characterization. The previous 2D representations can be unfolded by plotting $s(r)$ against the product of the density and the sign of the second eigenvalue of $H(\rho, r)$. This is normally expressed as $sign(\lambda_2)\rho(r)$. The sign of λ_2 measures charge accumulation in the perpendicular plane of the interaction. The function $sign(\lambda_2)\rho(r)$ is able to classify interaction types: strongly attractive (e.g., hydrogen bonds) < 0, van der Waals $\simeq 0$ (irrespective of the sign) and steric clashes > 0.

3

SEMI-EMPIRICAL AND LINEAR-SCALING DFT METHODS TO CHARACTERIZE DUPLEX DNA AND G-QUADRUPLEX DNA IN PRESENCE OF INTERACTING SMALL MOLECULES

3.1 INTRODUCTION

The interaction of ligands with DNA is a vital research subject with critical therapeutic consequences. Small molecules have shown enough anti-tumoral activity [169] binding efficiently to dDNA or GQ. For instance, cisplatin [203] is an effective drug which effectively binds to dDNA. However, the toxicity [42, 51], resistance and non-specific interactions of available drugs make desirable the quest for new molecules that target DNA, in its canonical dDNA form or in other secondary DNA structures, such as GQ. In this sense, theoretical methods that help to elucidate the nature of the interactions between ligands and DNA are of paramount importance. Studies on the interaction of ligands with dDNA structures are available in the literature [22, 59, 94, 171]. However, studies with other DNA structures, such as GQ are scarcer [6, 269]. GQ have raised considerable interest during the past years for the development of therapies against cancer. These non-canonical structures of DNA may be found in telomeres and/or oncogene promoters, and it has been observed that the stabilization of such GQ may disturb tumor cell growth [45, 160].

There are different theoretical methods available for the study of DNA and the state-of-the-art for the GQ modeling has been reviewed very recently in our team [154]. These methods range from MD simulations with classical Force Fields (FF) to accurate QM calculations

to understand specific local interactions in detail. The use of classical FF, such as OL15 or bsc [115, 268], for the study of DNA has demonstrated high precision and reliability[74] but they show limitations to deal with GQ since their parameters are usually optimized considering dDNA structures and not GQ. In this sense, optimization of FF parameters for the correct description of the GQ has been the subject of work by Sponer et al [268].

Despite the complexity of biological systems, many phenomena may be studied just taking into account a relatively localized region. Assuming that the QM treatment is needed, the choice is between a QM cluster approach or the use of the hybrid QM/MM approach [19]. Although the QM/MM approach has been extensively used to date [153, 157], it also presents shortcomings that may affect the quality of the results: *a*) The QM/MM method is based on partitioning the system of interest into a QM and a MM region, which introduces some arbitrariness in the calculation; *b*) the partition can lead to the cleavage of covalent bonds, and one needs an appropriate treatment of this boundary; *c*) electrostatic interactions between the QM and MM regions may be considered at different levels. If one needs to overcome the inherent limitations of an artificial QM/MM partition or the property of interest cannot be localized in a small region, a full SE description of the system may be a good alternative [49, 107]. SE methods fill the gap between MM and first-principles QM methods, being few orders of magnitude slower than MM methods but still orders of magnitude faster than first-principles QM methods. On the other hand, SE methods may provide reasonable accuracy in geometries and energies, although this is often system-dependent. In addition, SE methods can take polarization and charge transfer effects into account but they have usually problems describing dispersion and hydrogen bonds [58]. Nevertheless, more recently, corrections have been added to overcome these difficulties to improve their performance considerably [69, 134, 245, 246, 262]. That is, corrections were added by Rezac et al. in 2009 in one of the most pop-

ular Hamiltonians, PM6, to improve the description of non-covalent interactions in the modified PM6-DH Hamiltonian [196], which later was revised in the PM6-DH2 Hamiltonian by Korth et al. [133, 264, 265]. Recently, in 2016, an extensive review of the SE methods for non-covalent biochemical interactions have been carried out by Qiang Cui et al. [49].

The software improvements on theoretical methods and their optimization for parallelization make possible nowadays accurate *ab initio* calculations, in which the computational cost scales linearly with the system size, such as LS-DFT [40] or near linear-scaling Coupled-Cluster [199]. The treatment of the entire system with a first-principles quantum method ensures an accurate description of the interaction between ligands and biomolecules. Therefore, it can be of high relevance to study DNA-Ligand interactions.

In the present chapter, we use a LS-DFT approach, namely, with the SIESTA [229] code, using the LMKLL density functional [144], and the DZDP basis set [144], to analyze the interaction of small chemical species with dDNA and GQ. The results are compared with the SE methods PM6-DH2 [133, 196] and PM7 [235] (using the MOPAC package [235, 236]), QM/MM calculations (by using Gaussian [73]), near linear-scaling DLPNO-CCSD(T) calculations [200] (implemented in the ORCA *v4.2.1* package [167]), and benchmark database [210]. Modeling of GQ have been the subject of substantial interest recently [154]. As far as we know, this is the first time that LS-DFT methods have been used to study ligand-GQ interactions by taking into account the whole non-canonical secondary structure found in the PDB of more than 1000 atoms, which includes a GQ structure, without reduction to smaller models. We also demonstrate the adequacy of these LS-DFT approaches to treat DNA in its multiple forms and their interactions with small chemical species.

3.2 METHODS

To perform SE geometry optimizations, we used MOPAC *v8.0.0* with the PM6-DH2 and PM7 Hamiltonians. The default Eigenvector Following routine (EF) was used [17], whereas the default SCF criterion was changed to $1x10^{-8}$. The minimum trust radius was set at 0.0001 Å/rad and a damping factor of 10 was added with the SHIFT method to improve the SCF procedure [162]. For the LS-DFT calculations, SIESTA *4.1-b3* software was used [229]. Geometry optimizations were performed with the LMKLL van der Waals functional [144], which includes dispersion corrections, being highly appropriate for the characterization of weak forces within the ligand-DNA interaction. The modified Broyden algorithm was used for geometry optimization [117]. SCF convergence was accelerated with the Pulay method [21] keeping a history of 4 past density matrices, the density matrix mixing weight was set to 0.005. For the basis set, a 30 meV energy shift was used along with a 150 Ry mesh cut off for real space integration. We also used a SCF tolerance of $1x10^{-5}$ eV and we set the max force tolerance at 0.02 eV/Å for the dDNA and at 0.1 eV/Å in the case of GQ structures. In the case of GQ systems the max force tolerance was established after doing some tests where it was observed that although the tolerance is quite high, it does not compromise either the total energy or the relaxed geometry, and at the same time it allows to reduce the computation time (see Table 1). An optimized double-zeta plus double polarization (DZDP) basis set was used for each atom [119] along with Troullier-Martins norm-conserving pseudopotentials [129, 249] that were generated with the ATOM package included in SIESTA software, whereas for the G-tetrads used for energetic calibration, the psml pseudopotentials [76] from the website www.pseudo-dojo.org were used, after verifying that they give the same results as the Troullier-Martins norm-conserving pseudopotentials (See Appendix A. It must be mentioned that, in the

Table 1: Total Energy (eV), Wall Time (s) and RMSD (Å) for the G_4MG_4 system [266] geometry optimization with different max force tolerance. RMSD value was calculated taking as reference structure the geometry of the literature.

Max Force Tolerance	Total Energy	Wall Time	RMSD
0.5	-4659.90	19171.3	0.01
0.2	-4659.97	19287.7	0.02
0.1	-4659.99	21431.6	0.02
0.07	-4660.00	23733.8	0.03
0.05	-4660.00	31654.6	0.06
0.02	-4660.01	40450.2	0.07

SIESTA method the matrix elements are computed with linear-scaling algorithms, while the diagonalization is proportional to $O(N^3)$.

DLPNO-CCSD(T) [200] single-point calculations were also performed with the ORCA *4.2.1* software [167]. The Ahlrichs's def2-SVP basis set [257] was used for all the calculations with the corresponding auxiliary bases of Weigend for RI-J [256] and RIJCOSX approximations [136].

We also ran QM/MM geometry optimizations for the 1N37 and 2JWQ PDB structures, that were performed at M11L/6-31+G(d,p):AMBER and B3LYP-D3(GD3BJ)/6-31+G(d,p):AMBER level, respectively, as implemented in Gaussian*16* [73]. Quadratic Convergence (QC) [12] SCF procedure was used with a maximum amount of 1500 cycles and the maximum size for an optimization step was changed from 30 to 1. In both cases the containing ligand was treated as the QM part and the rest of the molecule (the DNA) as the MM layer, no boundaries needed to be used, because the ligand and DNA are not linked covalently.

3.2.1 *Considered structures*

3.2.1.1 *DNA base pairs.*

Initial geometries were taken from the benchmark database performed by Hobza and coworkers, which contains computationally optimized and experimental DNA base pair structures interacting through H-bonds and by $\pi - \pi$ stacking [120]. Interaction energies (ΔE_{int}) for these systems were computed by single-point calculations, at the different levels we compare in this study, and subtracting from the total energy (E_{tot}) the energy of each DNA base fragment (E_{frag}) as shown in Eq 3.1:

$$\Delta E_{int} = E_{tot} - E_{frag1} - E_{frag2} \tag{3.1}$$

3.2.1.2 *Intercalated Phenantholine in DNA base pairs.*

Optimized structures at M06-2X/6-31+G(d,p) level were taken from a previous work [78] and single-point calculations were performed at the different levels of calculation we compare in this work. The interaction energy for the phen/DNA systems was calculated as shown in Eq 3.2, where one fragment is composed of all DNA atoms and the second is the phenanthroline (phen) ligand:

$$\Delta E_{int} = E_{tot} - E_{DNA} - E_{phen} \tag{3.2}$$

3.2.1.3 *G-tetrads.*

We considered four model systems from the work of Fonseca Guerra and coworkers: G_4MG_4, aG_4MG_4, GQM, and $GQ_{4Na}M$, where M corresponds to the different metal cations (Li, Na, K, Rb, and Cs) placed in the ion channel [266] The first two models (G_4MG_4 and aG_4MG_4) only

contain the metal cation and the guanine bases, whereas the other two models (GQM and $GQ_{4Na}M$) contain also the sugar and phosphate backbone or side loop, which is terminated by H^+ or Na^+, respectively, to compensate for the negative charge of the phosphate group [172]. For these systems, single-point calculations were performed at the different levels of calculation we compare in this work and the interaction energy was calculated following the same formula as in the original work, Eq 3.3:

$$\Delta E_{int} = E(G_4MG_4) - E(G_4G_4) - E(M) \tag{3.3}$$

where G_4MG_4 is replaced by aG_4MG_4, GQM, or $GQ_{4Na}M$, and the empty scaffold G_4G_4 is replaced by aG_4G_4, GQ, or GQ_{4Na}, which are the same structures without the metal cation.

3.2.1.4 *1N37 and 2JWQ systems*

To perform larger geometry optimizations, two different DNA structures were taken from the PDB: 1) the 1N37 octamer [218], which contains Respinomycin D ligand intercalated in a dDNA structure; and 2) the GQ structure 2JWQ [108], which has two MMQ-1 units, each bounded at the end-staking of both sides of the GQ. The systems were neutralized by adding an alkaline cation close to each phosphate group at a distance of 2.8 Å, Na^+ in the dDNA and K^+ for the GQ. In the case of the GQ, two additional K^+ were added, centered in the ion-channel between G-tetrads, resulting in a +2 charged system.

3.3 RESULTS AND DISCUSSION

3.3.1 *Geometrical discussion*

3.3.1.1 *Duplex DNA*

Geometry optimizations were run departing from the PDB structures with SE, QM/MM, and LS-DFT methods. In Figure 7, we plot the superposition of the 1N37 structure from the PDB with the optimized geometries. SE results are the ones that show the highest deviation from the initial reference structure giving the highest Root-Mean-Square Deviation (RMSD) values. Namely, PM6-DH2 shows a RMSD of 1.84 Å, although the structure seems to be maintained. However, PM7 gives a much higher RMSD, 3.02 Å, and the general structure is clearly not well superimposed. The QM/MM method gives a slightly better result than PM6-DH2 with a 1.78 Å RMSD value and a lower deviation for the structure corresponding to the ligand than for the dDNA part. Finally, the LS-DFT method is by far the best method reproducing the PDB structure, giving an excellent RMSD value of 0.45 Å. Such results show that the SIESTA method and software not only has good performance in terms of computing time for large systems of thousands of atoms, but also that the obtained structures are very accurate in terms of geometry.

To further calibrate the performance of these methods for structural characterization, we also used other geometrical parameters characteristic of these systems with drugs intercalating between DNA base pairs. We analyzed the hydrogen bond lengths, the so-called twist angle (ϑ) and the rise (R) parameter, and we compared them to the counterparts of the original 1N37 PDB structure. Such values are depicted in Table 2.

The used nomenclature for each atoms can be seen in Figure 8. We plot also a dashed line joining the C_8 atom of the purine to the C_6

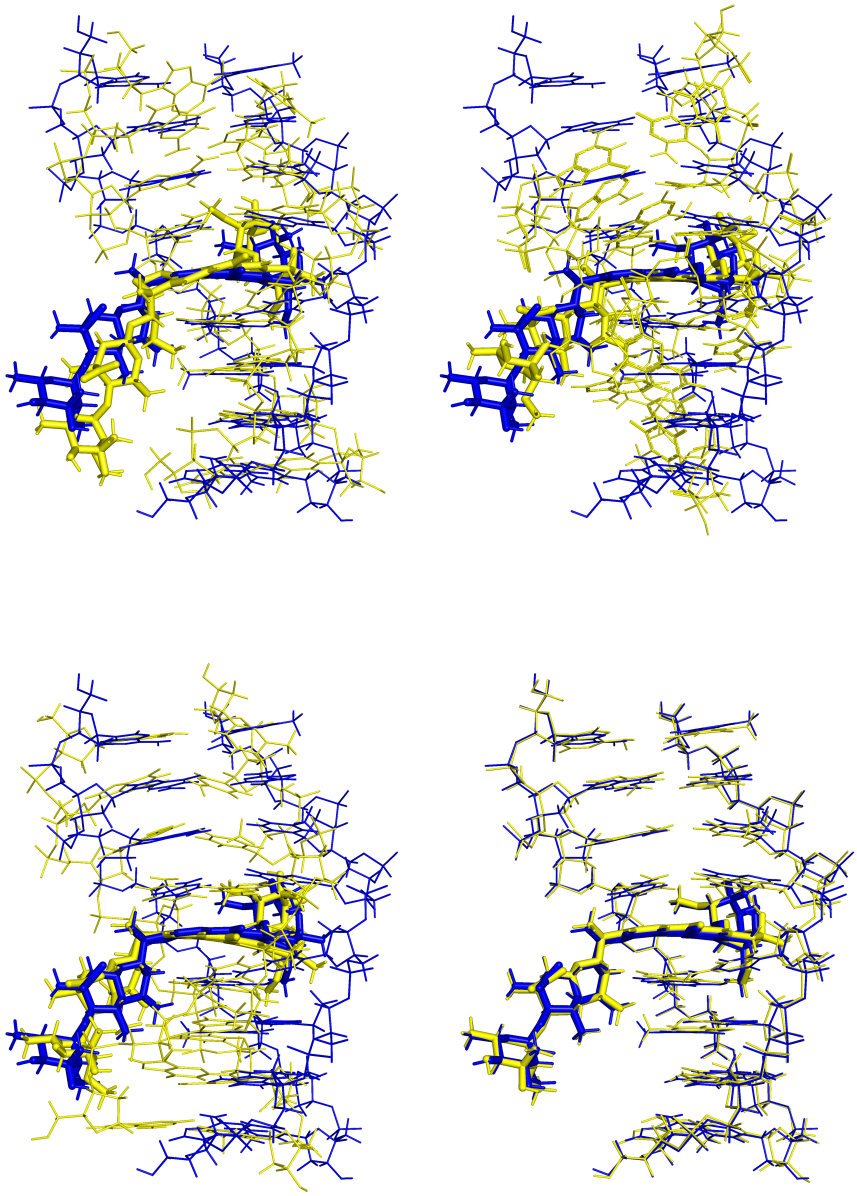

Figure 7: Overlap structures of 1N37 (in blue) with the optimized geometries (in yellow). From left to right, and top to bottom with the methods PM6-DH2, PM7, M11L/6-31+G(d,p):AMBER and LMKLL/DZDP. RMSD values of 1.84 Å, 3.02 Å, 1.78 Å and 0.45 Å were obtained, respectively.

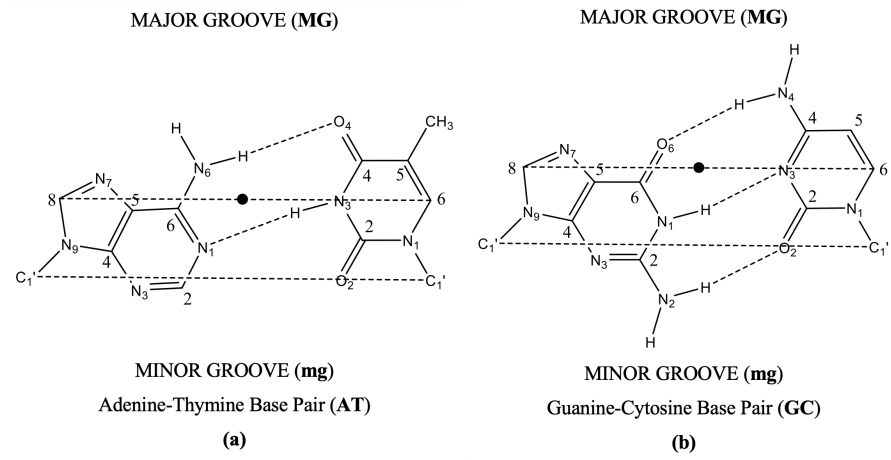

Figure 8: Scheme of the base pairs AT (a) and GC (b). The dashed line C_6–C_8 represents the long base-pair axis, which is roughly parallel to the C'_1–C'_1 line, where C'_1 stands for the sugar carbon atoms bonded to the bases. The twist angle (ϑ) is defined as the rotation around the midpoint of the C_6–C_8 axis (denoted by a dot).

atom of the pyridine. The twist angle (ϑ) is defined as the rotation of one base pair around the center of its C_6–C_8 axis.

First, it is worth mentioning that the PM7 SE method gives the worst geometrical results for the 1N37 system. Although the main structural features are kept, some of the base pairs, such as A_3-T_6 or T_8-A_1, are separated. The structure is not kept straight, in a conformation with all bases stacked, but curved in a "C" shape. For these reasons, we will not further consider this structure for deeper geometrical analysis.

Let us start analyzing the twist angle parameter, calculated in the same way as in previous works [75]. LMKLL is the method that gives the best results for the twist angle, with differences from 0.0° to 1.3° in most cases. Only two twist angles differ by more than 2°: the twist angle formed between C_4-G_5 and G_5-C_4 and that formed between G_5-C_4 and T_6-A_3. On the other hand, it is unclear which approach gives the worst results for the twist angle. In some cases, QM/MM yields the

Table 2: Hydrogen bond lengths, rise (R) parameter and twist angle (ϑ) of the original 1N37 PDB structure in parentheses, after LS-DFT optimization at LMKLL/DZDP level with SIESTA in bold, after optimization with PM6-DH2 in italics, after optimization with PM7 in italics plus bold and finally, QM/MM optimization at M11-L/6-31+G(d,p):AMBER level in normal script (see Figure 8 to check the labels of the atoms corresponding to the hydrogen bonds).

Base Pairs	Purine···Pyrimidine	Distance (Å)	R (Å)	ϑ(°)
A_1-T_8	N_6···O_4	(2.93)/**3.16**/*2.78*/***2.98***/2.88		
	N_1···N_3	(2.84)/**2.85**/*3.12*/**3.15**/2.92		
	C_2···O_2	(3.47)/**3.38**/*4.38*/**3.97**/3.66		
			(2.85)/**3.61**/*3.44*/**3.73**/3.53	(23.9)/**23.2**/*27.2*/***8.0***/22.2
G_2-C_7	O_6···N_4	(2.83)/**2.92**/*2.87*/***3.02***/2.89		
	N_1···N_3	(2.90)/**2.95**/*2.86*/**2.93**/2.90		
	N_2···O_2	(2.84)/**2.93**/*2.83*/**2.83**/2.78		
			(3.60)/**3.15**/*3.47*/***1.42***/2.90	(23.3)/**21.8**/*26.2*/***-43.4***/30.5
A_3-T_6	N_6···O_4	(2.87)/**3.11**/*2.85*/***5.61***/3.02		
	N_1···N_3	(2.94)/**3.00**/*3.00*/**7.33**/2.87		
	C_2···O_2	(3.70)/**3.84**/*4.00*/***8.84***/3.45		
			(3.54)/**3.46**/*3.54*/***2.89***/3.57	(21.6)/**21.6**/*20.5*/***15.3***/20.9
C_4-G_5	O_6···N_4	(2.92)/**3.80**/*2.87*/***2.89***/2.95		
	N_1···N_3	(2.91)/**3.29**/*2.84*/***3.00***/2.91		
	N_2···O_2	(2.80)/**2.87**/*2.82*/**2.92**/2.78		
			(5.66)/**5.75**/*5.94*/**5.51**/6.05	(5.2)/**24.5**/*30.5*/***15.4***/21.9
G_5-C_4	O_6···N_4	(2.87)/**2.93**/*2.85*/***2.88***/2.84		
	N_1···N_3	(2.88)/**2.99**/*2.87*/***2.91***/2.90		
	N_2···O_2	(2.83)/**2.96**/*2.83*/***2,86***/2.84		
			(3.52)/**3.31**/*3.16*/***3.10***/3.35	(25.0)/**22.7**/*26.2*/***28.4***/23.6
T_6-A_3	N_6···O_4	(2.96)/**3.27**/*2.84*/**3.25**/2.89		
	N_1···N_3	(2.89)/**3.04**/*3.05*/**2.89**/2.86		
	C_2···O_2	(3.61)/**3.79**/*4.18*/**3.35**/3.46		
			(3.15)/**3.07**/*3.31*/***2.90***/4.02	(26.8)/**26.5**/*26.2*/***20.2***/29.7
C_7-G_2	O_6···N_4	(2.80)/**2.88**/*2.83*/**2.94**/2.87		
	N_1···N_3	(2.94)/**3.01**/*2.85*/**2.98**/2.85		
	N_2···O_2	(2.99)/**3.08**/*2.86*/**2.97**/2.89		
			(3.18)/**3.14**/*2.73*/***0.44***/2.79	(0.7)/**-0.9**/*-12.5*/***18.8***/-17.0
T_8-A_1	N_6···O_4	(7.40)/**7.86**/*3.04*/**5.41**/3.46		
	N_1···N_3	(7.31)/**7.89**/*4.67*/***6.08***/7.93		
	C_2···O_2	(6.12)/**6.68**/*7.37*/**7.57**/11.19		

worst results, whereas in others, PM6-DH2 does. In general, the results for the twist angle given by the three different approaches agree with the analysis of the RMSD.

Regarding the Rise parameter (R), the QM/MM approach at M11-L/6-31+G(d,p):AMBER level of theory gives the worst results with differences from the PDB structure higher than 0.5 Å in some cases. On the other hand, as a general trend, the best results are again given by the LS-DFT. Nevertheless, the SE approach also gives excellent results for the rise parameter (R), in some cases (A_1-T_8/G_2-C_7, G_2-C_7/A_3-T_6, and A_3-T_6/C_4-G_5), even better than the LS-DFT approach. Thus, it is not clear which of the two approaches, LS-DFT or SE, including dispersion yield the best results.

Interesting trends are also observed for hydrogen bond distances. In the case of the hydrogen bond distances of the A-T base pairs, we see that the SE approach gives the worst behavior. Indeed, there is a considerable lengthening, especially for the $N_1 \cdots H\text{-}N_3$ and $C_2\text{-}H \cdots O_2$ hydrogen bonds. LS-DFT approach gives the best results for the hydrogen bond distances again, although in general, they are elongated with respect to the PDB structure, especially for the $N_1 \cdots H\text{-}N_3$ and $C_2\text{-}H \cdots O_2$ hydrogen bonds. Nevertheless, it must be mentioned that the QM/MM approach at M11-L/6-31+G(d,p):AMBER also gives the best results in some of the $N_6\text{-}H \cdots O_4$ hydrogen bonds. On the other hand, there is a better agreement between theoretical hydrogen bond distances and those corresponding to the original PDB structure for the G-C base pairs, with LS-DFT giving the best results.

Summarizing, the SE approach with the PM6-DH2 Hamiltonian gives in some cases better results than the most popular QM/MM approach at M11-L/6-31+G(d,p):AMBER level of theory. On the other hand, the LS-DFT method gives an excellent agreement between the geometrical parameters of the PDB structure and those of the optimized structure. SIESTA software gives us excellent results for this biomolecular system of 665 atoms in a reasonable computation time.

3.3.1.2 *G-quadruplex*

We carried out an analysis for the systems based on GQ by using similar structural parameters as for the dDNA (See Table 3): RMSD from the PDB reference structure, hydrogen bond lengths, rise distances, and twist angle, in this case as defined by Chung et al. [53], where the xy plane is defined by three guanine O_6 of the same G-tetrad. Then, the R mean distance between the two G-tetrads is defined as the difference between the mean z value of the atoms of the upper G-tetrad and the one of the atoms of the lower G-tetrad. The ϑ twist angle, is defined as the angle between the lines formed by the guanine C_8 and the midpoint between N_1 and C_2, as can be observed in Figure 9. See also Figure 10 to see the used nomenclature for the GQ system

The 2JWQ PDB structure for the GQ was optimized at PM6-DH2 and PM7 level (in the case of the SE approach), at B3LYP-D3(GD3BJ)/6-31+G(d,p):AMBER level (in the case of the QM/MM approach) and at LMKLL/DZDP level (an LS-DFT method), see Figure 11. SE methods yielded higher RMSD values than the LS-DFT method. The PM6-DH2 and PM7 methods gave RMSD values of 0.98 and 2.06 Å, respectively. Quite interestingly, the RMSD values obtained with the SE models for the systems based on GQ were lower than those obtained for dDNA, even though the GQ used in this work has approximately twice as many atoms as the studied dDNA systems. This better agreement could be due to a more rigid structure of the GQ because the stacking of the G-tetrads creates a more rigid structure, which in turn increase the rigidity of the sugar and phosphate backbone.

The LMKLL LS-DFT method gave the best agreement to the experimental structure again, with a RMSD of 0.24 Å. On the other hand, QM/MM and PM6-DH2 show a similar RMSD, 1.05 Å for the former and 0.98 for the latter, whereas PM7 gives the worst result, namely, a RMSD of 2.06 Å. There is a tendency for the different computational methods to reduce the range in distances between the heavy atoms in

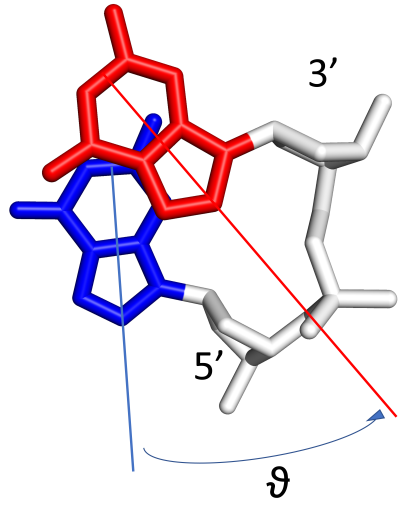

Figure 9: Representation of the stacking of two G-tetrad guanines. The ϑ twist angle, is defined as the angle between the lines formed by the guanine C_8 and the midpoint between N_2 and C_6.

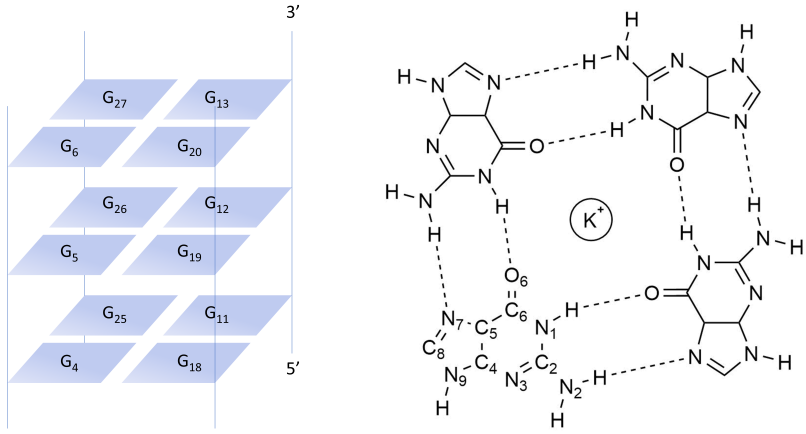

Figure 10: On the left side, scheme of the stacking of three G-tetrads of guanines composing the 2JWQ PDB structure, and nomenclature for each guanine base. On the right side, a G-tetrad, with the numbers used in the nomenclature.

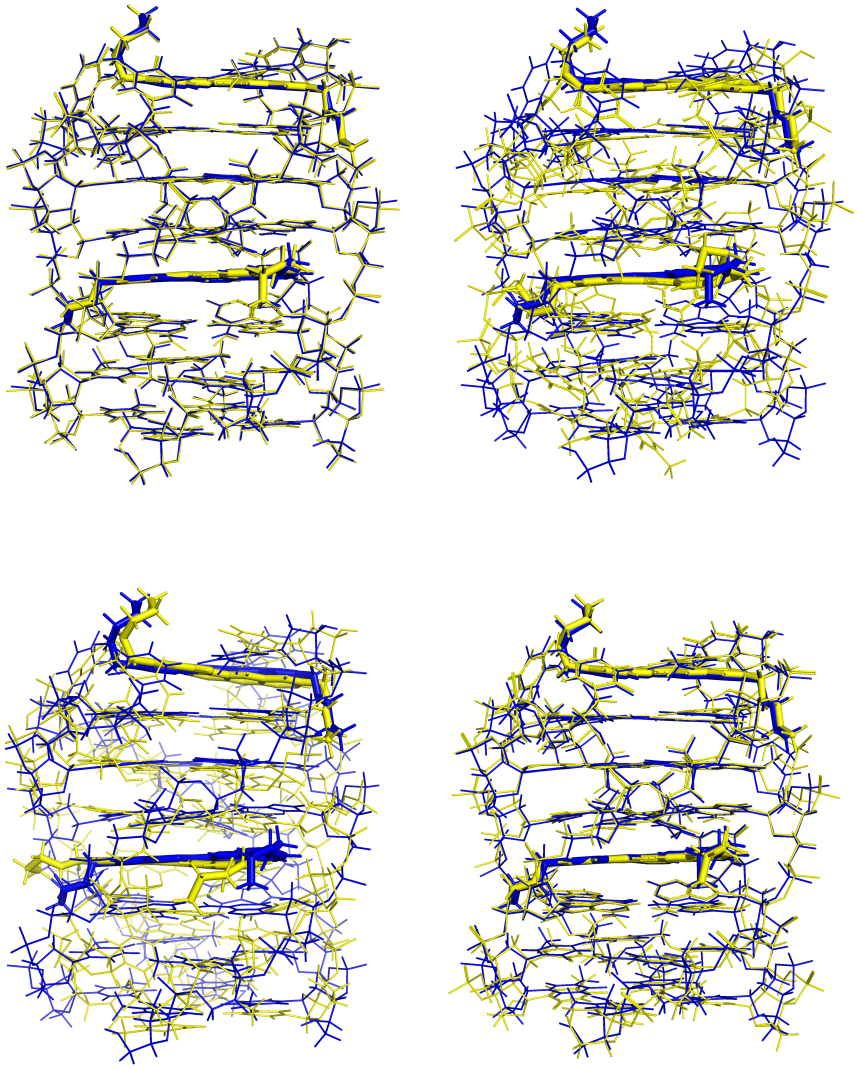

Figure 11: Superposition of the 2JWQ PDB structure (in blue) with the optimized geometries (in yellow) with the PM6-DH2, PM7, B3LYP-D3 (GD3BJ)/6-31+G(d,p):AMBER and LMKLL methods, from left to right and from top to bottom. RMSD values of 0.98 Å, 2.06 Å, 1.05 Å and 0.24 Å, respectively.

Table 3: Hydrogen bond lengths, rise (R) parameter and twist angle (ϑ) of the original 2JWQ PDB structure in parentheses, after LS-DFT optimization at LMKLL/DZDP level with SIESTA in bold, after optimization with PM6-DH2 in italics, after optimization with PM7 SE method in italic plus bold, and at the B3LYP-D3(GD3BJ)/6-31+G(d,p):AMBER level in normal script (see Figure 10 to check the labels of the atoms corresponding to the hydrogen bonds and the labels for guanines in the tetrads).

Tetrads	Guanine···Guanine	Distance (Å)	R (Å)	ϑ(º)
$G_{4,11,18,25}$	$G_4O_6\cdots G_{18}N_1$	(3.01)/**2.90**/*2.90*/***3.00***/2.92		
	$G_4N_7\cdots G_{18}N_2$	(2.88)/**3.00**/*3.04*/***3.01***/2.95		
	$G_{18}O_6\cdots G_{11}N_1$	(3.14)/**2.94**/*2.88*/***2.96***/2.94		
	$G_{18}N_7\cdots G_{11}N_2$	(2.77)/**2.93**/*3.00*/***3.07***/2.88		
	$G_{11}O_6\cdots G_{25}N_1$	(2.85)/**2.90**/*2.90*/***2.81***/2.89		
	$G_{11}N_7\cdots G_{25}N_2$	(3.41)/**3.16**/*3.06*/***3.15***/3.01		
	$G_{25}O_6\cdots G_4N_1$	(2.92)/**2.91**/*2.82*/***2.88***/2.89		
	$G_{25}N_7\cdots G_4N_2$	(2.93)/**3.02**/*2.97*/***3.06***/2.88		
			(3.38)/**3.41**/*3.31*/***3.33***/3.39	(28.7)/**25.3**/*26.8*/***25.3***/24.8
$G_{5,12,19,26}$	$G_5O_6\cdots G_{19}N_1$	(3.10)/**2.91**/*2.74*/***2.82***/3.17		
	$G_5N_7\cdots G_{19}N_2$	(3.45)/**2.98**/*3.20*/***2.96***/2.92		
	$G_{19}O_6\cdots G_{12}N_1$	(2.53)/**3.08**/*2.86*/***2.87***/2.98		
	$G_{19}N_7\cdots G_{12}N_2$	(2.43)/**2.98**/*3.14*/***3.01***/2.86		
	$G_{12}O_6\cdots G_{26}N_1$	(2.91)/**2.85**/*2.79*/***2.98***/2.95		
	$G_{12}N_7\cdots G_{26}N_2$	(3.77)/**3.38**/*3.53*/***2.96***/3.13		
	$G_{26}O_6\cdots G_5N_1$	(3.08)/**2.92**/*2.80*/***2.89***/2.90		
	$G_{26}N_7\cdots G_5N_2$	(3.17)/**3.00**/*3.06*/***2.89***/2.96		
			(3.19)/**3.26**/*3.17*/***3.21***/3.42	(16.4)/**21.9**/*19.5*/***30.3***/28.5
$G_{6,13,20,27}$	$G_6O_6\cdots G_{20}N_1$	(2.83)/**2.85**/*2.77*/***2.88***/2.75		
	$G_6N_7\cdots G_{20}N_2$	(2.85)/**3.02**/*3.06*/***2.93***/3.02		
	$G_{20}O_6\cdots G_{13}N_1$	(2.99)/**2.99**/*2.80*/***2.87***/2.82		
	$G_{20}N_7\cdots G_{13}N_2$	(2.80)/**3.00**/*3.18*/***3.05***/2.89		
	$G_{13}O_6\cdots G_{27}N_1$	(2.89)/**2.99**/*2.83*/***2.87***/2.87		
	$G_{13}N_7\cdots G_{27}N_2$	(3.05)/**3.04**/*3.17*/***2.98***/2.84		
	$G_{27}O_6\cdots G_6N_1$	(2.86)/**2.86**/*2.80*/***2.93***/2.85		
	$G_{27}N_7\cdots G_6N_2$	(2.91)/**3.04**/*3.02*/***2.98***/2.86		

the hydrogen bonds with respect to the values observed in the PDB. The PDB structure shows a range in $O_6 \cdots N_1$ distances between 2.53-3.14 Å, whereas for LMKLL, QM/MM, PM6-DH2, and PM7, the ranges are 2.85-3.08 Å, 2.75-3.17 Å, 2.74-2.9 Å, and 2.81-3.00 Å, respectively. The differences are even higher in the case of the $N_7 \cdots N_2$ distances, in which the PDB range is 2.43-3.77 Å while for the LMKLL, QM/MM, PM6-DH2, and PM7 they are 2.93-3.38 Å, 2.82-3.13 Å, 2.97-3.53 Å and 2.89-3.15 Å, respectively. The LMKLL is again the method that better reproduces these distances, followed by QM/MM, PM6-DH2, and PM7 methods. These changes in distance, although small, cause the guanine bases to rotate directly affecting the twist angle. To calculate the twist angle we used a method devised by Phan et al. [53] as said above. The rotation angle between the G-tetrads was established as the average value obtained by calculating the angle of two guanines stacked between the vectors formed for each guanine by the coordinate of the C_8 atom and the midpoint of the coordinates of the N_1 and C_2 atoms. We have found that the twist angle for $G_{4,11,18,25}$ - $G_{5,12,19,26}$ is well reproduced by all methods, with a maximum deviation from the original PDB (28.7°) of 3.9° in the case of PM7. However, the twist angle for $G_{5,12,19,26}$ - $G_{6,13,20,27}$ (16.4°) is reproduced appropriately by LMKLL (21.9°) and QM/MM (19.5°), whereas PM6-DH2 (30.3°) and PM7 (28.5°) SE methods show larger discrepancies. Finally, the Rise distance is well reproduced by all methods with differences with respect to the PDB reference structure within 0.2 Å in all cases.

In summary, the LS-DFT at the LMKLL/DZDP level of theory is the method that better describes the geometrical structure of the GQ system. In general, we observe a better agreement between the optimized and PDB structures for GQ than for dDNA, which can be related to the more rigid structure of the former. The geometrical results obtained for 1N37 dDNA and 2JWQ GQ confirm that our LS-DFT approach is appropriate to optimize DNA-type biomolecules. On the other hand, PM6-DH2 describes reasonably well both dDNA and GQ systems but not the

PM7 method. In addition, the PM6-DH2 approach may yield geometries of similar quality as for the popular QM/MM methodology. However, the PM7 Hamiltonian shows the most significant deviations in geometries. Finally, it must be said that the popular QM/MM approach used in this work, namely, at B3LYP-D3(GD3BJ)/6-31+G(d,p):AMBER, may also lead to qualitatively correct structures.

3.3.2 *Energetics discussion*

3.3.2.1 *DNA base pairs*

Figure 12 shows graphically the trends for the interaction energies (ΔE_{int}) of different DNA hydrogen-bonded and stacked base-pairs in the gas phase for the different computational methods studied in our work. The corresponding values are depicted in Table 4.

It is observed that the DLPNO-CCSD(T) and LMKLL methods reproduced the interaction energies of the benchmark references [120] accurately. The obtained ΔE_{int} for LMKLL and DLPNO-CCSD(T) are virtually identical, and both methods reproduce the reference interaction energies. The correlation coefficient with respect to the reference data is 0.996 for both methods, denoting a very good performance to describe the trends in interaction energies of these systems. The mean absolute error (MAE) for DLPNO-CCSD(T) is 2.1/1.8 kcal mol^{-1} for hydrogen-bonded/stacked base pairs, whereas the MAE for LMKLL is only 1.1/2.1 kcal mol^{-1}, respectively. In the case of systems where H-bonding links the bases, the interaction energies obtained by LS-DFT are systematically higher in absolute value than those obtained with DLPNO-CCSD(T). In contrast, for the stacked base pair systems, the ΔE_{int} are similar in all cases. On the other hand, the LS-DFT H-bonded systems are slightly more accurately described than the stacked ones. This fact could be due to the use of the LMKLL functional, which improves the description of non-covalent interactions, but especially when hy-

Table 4: Interaction energies (kcal mol^{-1}) for the different studied DNA base pairs. Abbreviations used in the first column: A, T, C, G – adenine, thymine, cytosine, guanine; m – methyl-; WC, HB – Watson-Crick, Hoogsteen: OG, EG – optimized geometry, experimental geometry.

	System	Reference	DLPNO-CCSD(T)	LMKLL	PM6-DH2	PM7
H-bonded base pairs	G-C WC (OG)	-32.06	-33.42	-32.18	-18.06	-16.29
	mG-mC WC (OG)	-31.59	-33.53	-32.09	-17.78	-16.02
	A-T WC (OG)	-16.86	-18.43	-17.97	-8.60	-6.86
	mA-mT WC (OG)	-18.16	-19.54	-18.84	-9.09	-6.26
	A-T WC (EG)	-16.40	-18.32	-17.95	-8.46	-6.81
	G-C WC * (EG)	-35.80	-36.91	-34.84	-17.85	-16.03
	A-T WC (EG)	-18.40	-20.38	-19.93	-8.37	-6.99
	G-A HB (EG)	-11.30	-14.80	-13.67	-5.83	-6.16
	C-G WC (EG)	-30.70	-33.96	-32.11	-18.04	-16.45
	G-C WC (EG)	-31.40	-34.19	-32.02	-18.06	-16.47
	MAE	–	2.08	1.08	9.15	11.61
Stacked base pairs	G-C (OG)	-19.02	-21.88	-21.03	-10.99	-10.76
	mG-mC (OG)	-20.35	-23.54	-21.72	-10.28	-9.58
	A-T (OG)	-12.30	-14.66	-14.25	-5.22	-4.00
	mA-mT (OG)	-14.57	-17.52	-17.37	-6.14	-4.65
	A-T (EG)	-8.10	-9.37	-11.89	-3.13	-3.50
	G-C (EG)	-7.90	-8.13	-8.84	-8.00	-2.90
	A-C (EG)	-6.70	-8.33	-9.54	-2.43	-0.51
	T-G (EG)	-6.20	-7.81	-9.87	-3.52	-3.91
	C-G (EG)	-7.70	-8.54	-9.03	-6.51	-8.80
	A-G (EG)	-6.50	-9.33	-8.80	-6.58	-4.49
	C-G (EG)	-12.40	-12.71	-13.11	-18.06	-10.44
	G-C (EG)	-11.60	-12.74	-12.96	-10.84	-16.40
	MAE	–	1.77	2.09	6.12	7.37

*The geometries of both GC WC (EG) pairs are identical.

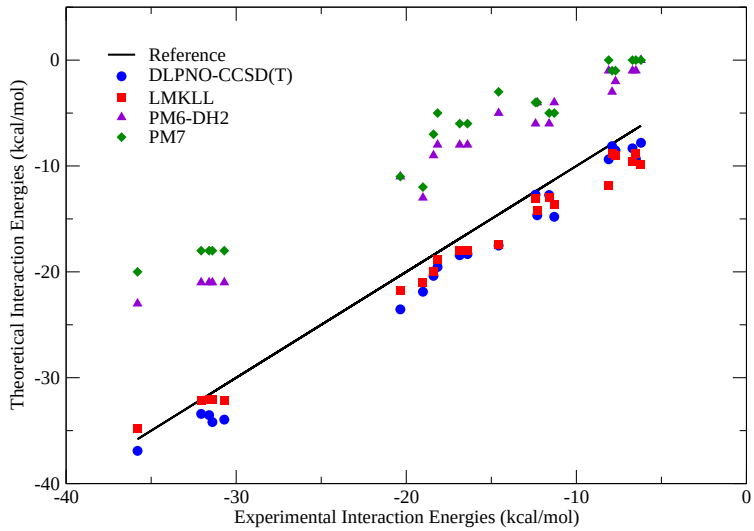

Figure 12: Interaction energies (kcal mol^{-1}) for DNA base pair systems, with CCSD energies from bibliographic reference [120] as the reference data. The correlation coefficients for the DLPNO-CCSD(T), LMKLL, PM6-DH2, and PM7 are 0.996, 0.996, 0.990, 0.984, respectively.

drogen bonding plays an important role [144]. On the other hand, the SE methods considerably underestimate the ΔE_{int} for these systems, with a MAE of 9.2 and 11.6 kcal mol^{-1} for PM6-DH2 and PM7, respectively, in the case of H-bonded base pairs and 6.1 kcal mol^{-1} and 7.4 kcal mol^{-1}, respectively, for stacked base pairs. Nevertheless, all methods show a similar r^2 correlation with respect to the reference data, indicating that even SE models may describe qualitatively the trends in interaction energies.

3.3.2.2 *Intercalated Phenantholine between DNA base pairs (phen/DNA)*

To validate the adequacy of these methods to treat the interaction of ligands intercalated between DNA base pairs through weak interactions, we used a previously characterized system [78], where a phen ligand is intercalated between two pairs of bases through both Major Groove (MG) and minor groove (mg). In this previous work, the MP2/6-31G*(0.25) theory level was used, based on the correction performed by Reha et al., which considers a modification of d-polarized basis function [194], leading to a better agreement with CCSD(T) benchmark energies. We performed calculations for the same systems in this work, taking the DLPNO-CCSD(T) method as a reference. We computed the ΔE_{int} between the ligand (phen) and the DNA base pairs. This ΔE_{int} was calculated by subtracting from the total energy the energy of the separated fragments: phen ligand and the DNA fragment. The trends obtained for the interaction energies are presented in Figure 13, whereas the values are depicted in Table 5.

Table 5: Interaction energies (kcal mol^{-1}) for the stacked base pairs with the intercalated phen ligand. A-T/phen/T-A MG and A-T/phen/T-A mg corresponds to Adenine-Thymine base pair system with intercalated phen through the MG and mg, while G-C/phen/C-G MG and G-C/phen/C-G mg corresponds to Guanine-Cytosine base pair system with phen intercalated through the MG and mg.

System	DLPNO-CCSD(T)	LMKLL	PM6-DH2	PM7
A-T/phen/T-A MG	-37.53	-38.26	-7.58	-3.97
A-T/phen/T-A mg	-33.81	-36.39	-5.71	-1.70
G-C/phen/C-G MG	-42.06	-41.69	-11.91	-8.03
G-C/phen/C-G mg	-35.87	-35.80	-6.49	-2.31
MAE		0.94	29.40	33.32

The agreement between LS-DFT and the reference DLPNO-CCSD(T) results is outstanding. The difference in interaction energies is lower than 1 kcal mol^{-1} in most of the cases, whereas the MAE is only

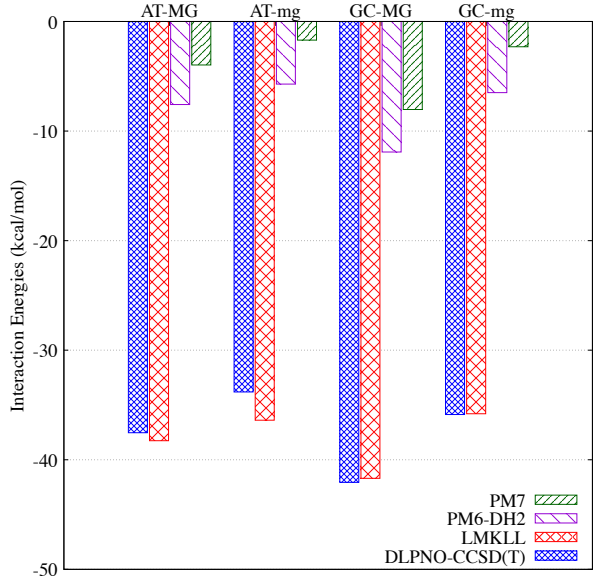

Figure 13: Interaction energies (kcal mol^{-1}) for the intercalated phen in DNA base pairs for the DLPNO-CCSD(T), LMKLL, PM6-DH2 and PM7 methods.

0.9 kcal mol^{-1}. It must be mentioned that the LMKLL functional has described with greater accuracy the ΔE_{int} of the systems containing guanine and cytosine. Moreover, the LMKLL functional describes the energetic trends very accurately and give more negative interaction energies for the MG structures than for the mg ones.

On the other hand, we observe that the SE methods show a poor performance with very small interaction energies, namely ΔE_{int}, which are around 30 kcal mol^{-1} smaller in absolute value than those of DLPNO-CCSD(T). The MAE for both methods is very significant, 29.4 and 33.3 kcal mol^{-1} for PM6-DH2 and PM7, respectively. Nevertheless, it must be said that the trends in energetics, in which the MG systems are more stable than the mg ones, were correctly described by SE methods. For both PM6-DH2 and PM7 methods, the G-C/phen/C-G structure intercalating via MG was described correctly as the most stable system with a ΔE_{int} of -11.9 and -8.0 kcal mol^{-1}, respectively.

Summarizing, the description of the intercalation of ligands with DNA through weak interactions requires a reliable method to describe non-covalent interactions. Our results point to LS-DFT with the LMKLL functional as a method with an excellent performance in ΔE_{int} for ligand-DNA systems.

3.3.2.3 *Structures based on G-tetrads*

The obtained interaction energies for the systems based on G-tetrads according to Eq. 3.3 are represented graphically in Figure 14 and their values depicted in Table 6 .

LS-DFT agrees not only with the DLPNO-CCSD(T) highly correlated benchmark calculations but also with the DFT calculations found in the bibliography for the same systems [266]. Our LS-DFT results tend to give smaller interaction energies than those found in the literature but the trend is described nicely with a correlation coefficient of 0.989 with respect to the benchmark DLPNO-CCSD(T) calculations. As in the

Table 6: Interaction energies (kcal mol^{-1}) for the different GQ structures (G_4MG_4, aG_4MG_4, GQM, and $GQ_{4Na}M$) with the different metal cations (Li, Na, K, Rb, and Cs) for the used computational methods (PM6-DH2, PM7, LMKLL/DZDP, and DLPNO-CCSD(T)/def2-SVP) along with the results found in the bibliography [266] at ZORA-BLYP-D3(BJ)/TZ2P level. The Mean Absolute Error (MAE) was calculated taking the DLPNO-CCSD(T)/def2-SVP energies as reference values.

	System	ZORA-BLYP-D3(BJ)	DLPNO-CCSD(T)	LMKLL	PM6-DH2	PM7
G_4MG_4	Li	-161.50	-153.67	–	-101.62	-117.22
	Na	-152.10	-149.87	-134.56	-127.28	-122.41
	K	-128.80	-129.86	-119.03	-78.80	-105.82
	Rb	-115.50	-115.54	-108.90	-67.96	-113.17
	Cs	-99.60	-97.20	-93.62	-97.89	-79.50
aG_4MG_4	Na	-145.80	-143.13	-129.79	-123.10	-118.35
	K	-126.60	-129.38	-118.00	-73.78	-102.97
	Rb	-114.70	-116.73	-108.84	-65.95	-102.83
	Cs	-99.20	-95.34	-93.30	-93.43	-78.05
GQM	Li	-165.70	-158.36	–	-102.40	-110.31
	Na	-156.60	-153.93	-137.09	-113.40	-113.72
	K	-134.70	-134.41	-123.74	-67.00	-107.12
	Rb	-119.10	-115.95	-115.17	-64.15	-117.80
	Cs	-104.40	-102.58	-101.09	-99.72	-94.30
$GQ_{4Na}M$	Na	-170.90	-170.04	-152.29	-115.60	-132.87
	K	-148.80	-148.73	-138.66	-88.68	-132.47
	Rb	-137.30	-136.45	-129.67	-77.25	-123.23
MAE		2.47	–	9.04	40.86	22.51

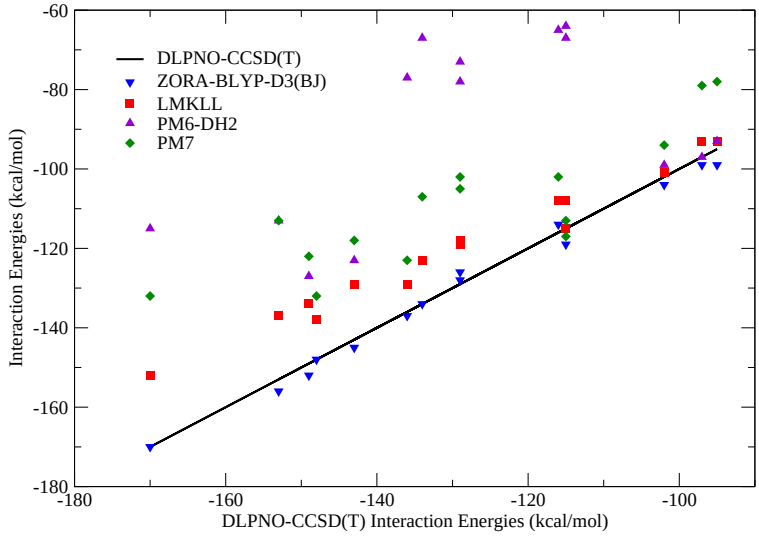

Figure 14: Interaction energies (kcal mol^{-1}) for the studied structures based on G-tetrads. r^2 correlation coefficients with respect to the DLPNO-CCSD(T) reference data for ZORA-BLYP-B3(BJ) [266], LMKLL, PM6-DH2, and PM7 are 0.993, 0.989, 0.509, 0.777, respectively.

case of the work of Fonseca-Guerra et al. [266], the Na metallic cation provides a significant stabilization for the G-tetrad structure and this stabilization decreases as the size of the metallic cation increases, (see Figure 15). On the other hand, the SE methods have a very poor behavior both qualitatively and quantitatively. There are significant MAEs for the interaction energies, namely 22.5 and 40.9 kcal mol^{-1} for PM7 and PM6-DH2, respectively. Moreover, the poor correlation coefficients, 0.509 for PM6-DH2 and 0.777 for PM7, reveal high limitations of SE methods to reproduce the correct trends in interaction energies. As observed in Figure 15, the SE methods show a less clear trend between stabilization of G-tetrads and metal size compared to those observed for DLPNO-CCSD(T) benchmark and LMKLL.

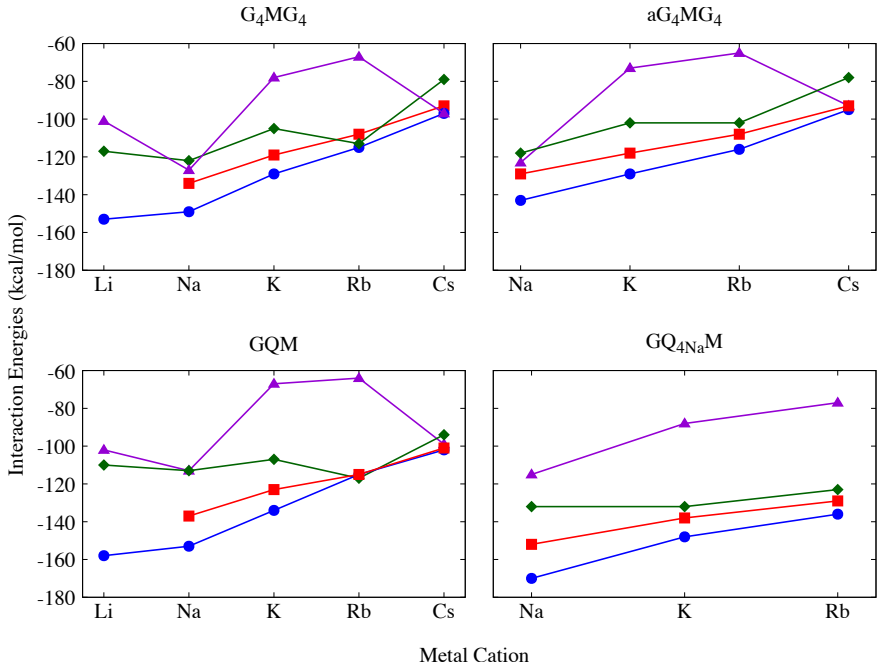

Figure 15: Interaction energies (kcal mol^{-1}) for the different G-tetrads studied systems (G_4MG_4 and aG_4MG_4 on top, and GQM and GQ_4NaM at bottom) with metal cations (Li, Na, K, Rb, and Cs) for each computational method used in this work. PM6-DH2 in purple, PM7 in green, LMKLL in red, and DLPNO-CCSD(T) in blue.

To sum up, considering all the interaction energies obtained for the different DNA systems, we can conclude that LS-DFT with the LMKLL functional can describe the weak interactions present in DNA systems. For G-tetrads, the ΔE_{int} deviate slightly from the energies found in the literature, but the energetic trends are in good agreement with those calculations. However, SE methods are very limited in the description of interaction energies. Therefore, the use of LS-DFT with the LMKLL functional including van der Waals corrections can be a very suitable strategy to analyze this kind of biological systems, in which a delicate balance of different non-covalent weak interactions is found.

3.4 CONCLUSIONS

Over the years, the computational power, the capacity for parallelization, and improvements in software have allowed the study of large biological systems using QM methods. In this sense, approaches based on LS-DFT are gaining in speed and efficiency, and they allow a more accurate description of the electronic structure of large biological systems. In this chapter we have studied three different DNA systems: i) DNA base-pairing, ii) models of dDNA interacting with the phen ligand and iii) G-tetrads stabilized with various alkaline metals. We have used a LS-DFT method (LMKLL/DZDP) and we have compared its performance with two SE methods incorporating dispersion corrections, PM6-DH2 and PM7, and QM/MM methods at B3LYP-D3(GD3BJ)/6-31+G(d,p):AMBER and at M11L/6-31+G(d,p):AMBER levels of theory. We have shown how the LMKLL applied through SIESTA reliably predicts both the geometries and the interaction energies for all these DNA systems, using experimental values and DLPNO-CCSD(T) benchmark calculations. On the other hand, the PM6-DH2 SE method has correctly described the geometries of these systems but not the PM7. However, both SE approaches are very limited in describing the interaction energies of these systems, with a degree of performance that

is system-dependent. The present work opens the door for the computational investigation of large DNA systems using LS-DFT methods, which is particularly interesting for GQ, in which a proper balance of different weak non-covalent interactions, metal-ligand interactions, polarization and charge transfer is needed to give a realistic description of the system.

4

UNRAVELING THE BINDING AFFINITY AND SELECTIVITY OF MOLYBDENUM (II) PHENANTHROLINE COMPLEXES WITH G-QUADRUPLEXES DNA. LINEAR-SCALING DFT STUDIES

4.1 INTRODUCTION

During the last decades, cancer has been one of the most studied diseases. Unfortunately, it has also been one of the diseases with lowest clinical success rates, where improving and finding new drugs against it is still a current subject of study [47]. One of the first effective drugs against cancer was the well-known cisplatin, reported by Rosenberg et al. in 1965 [202]. However, its toxicity limits its use in humans. A promising alternative is the use of non-covalent binding interacting ligands [27, 123, 142], where the interaction is strong enough to interrupt certain key biological processes such as DNA transcription or replication, which directly affect the growth of tumor cells.

On the other hand, an alternative and innovative strategy explored during the last years is the use of the non-canonical secondary DNA structures called G-quadruplexes GQ, which may be used as specific targets [18, 170]. GQ have been found at key points in the genome that directly relate it to cancer, such as telomeres or oncogene promoters [183, 224] and their stabilization is crucial for the disruption of the transcription and replication processes of DNA. In this sense, it must be said that even though there are different ways to stabilize GQ with ligands and other small molecules [32, 44, 141], the use of small

molecules not only stabilize GQ but also induces the formation of the GQ around the ligand [201].

Most of the studied molecules that can interact non-covalently with dDNA and GQ in a non-covalent way are planar organic molecules [6], such as phenanthroline and its derivatives [75, 77–79, 211].

In addition, many metal complexes, including phen and derivatives, have shown potential to inhibit the growth and survival of tumor cells, while being less toxic than cisplatin [43, 187, 220]. Alternatively, octahedral metal complexes containing planar ligands have also been investigated in a more reduced number of studies [100, 151, 152, 205, 223, 239, 263]. Among the used metallic elements, molybdenum complexes are very promising due to the fact that Mo is an essential trace in the human body, and it is characterized by its low toxicity [158]. Moreover, experimental studies demonstrated the efficacy of a molybdenum octahedral metal complex that contain phenanthroline [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] against different tumor cell lines [20].

GQ structures and its interaction with metal complexes have also been studied by computational methods [154]. MD simulations have been the most used computational methods to study the interaction of GQ structures with metal complexes. Different metal complexes have been studied such as various Salphen- and Schiff-base metal complexes [36, 68, 191, 261], although metal complexes containing phen derivatives have also been studied by means of MD [124]. GQ have also been studied by means of QM/MM [2, 102] and in the interaction with certain ligands [145, 163]. However, studies analyzing the interaction of such structures with metal complexes by QM/MM are scarce in the bibliography [86, 240]. As an alternative to MD and QM/MM, a recent work in the team performed a theoretical study on the interaction of this Mo- and phen-based metal complex with the dDNA using LS-DFT for the full system, defined as the complex and the entire DNA chain (a system with more than 500 atoms).

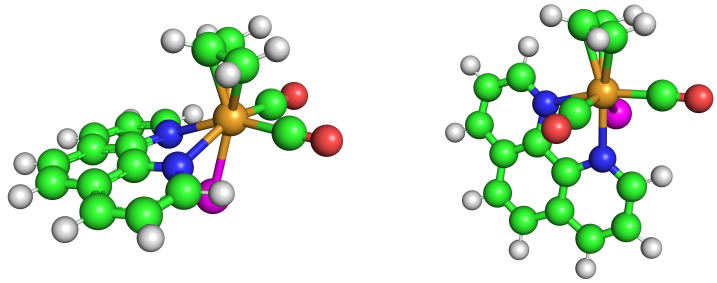

Figure 16: Equatorial (left) and Axial (right) isomers of the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex. (Mo: Orange, Br: Purple, O: Red, N: Blue, C: Green, H: White)

In this chapter, the non-covalent interaction of two isomers, Equatorial (Eq) and Axial (Ax), of the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] octahedral metal complex (Figure 16) with GQ is going to be studied using computational methods. LS-DFT will be used for the first time to study the interaction of GQ with metal complexes using the SIESTA software [7, 229], capable of performing DFT calculations for large biological systems of more than 1000 atoms. To study the nature of the interaction we will use tools like the QTAIM [122], NCI index [57, 118] and the EDA [106].

4.2 METHODS

Geometry optimizations were performed using a four stranded parallel GQ as starting point structure (PDBID: 2jwq) [108]. This structure has two quinacridine-based ligand (MMQ) units, each bounded by means of the end-staking mode of interaction at both sides of the GQ. In order to predict the binding of the two isomers of the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex with the 2jwq structure, molecular docking was carried out. The MMQ ligands were removed from the original structure of the PDB, which leaves a gap at the end-

stacking of the GQ where our [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex could potentially be introduced in the docking screening. The HEX *v8.0.0* package [166] was chosen for docking by using the protocol we describe in the following lines. In general, the manual recommendations were carried out for the DNA structure including the metal complex. Nevertheless, we changed some of the default parameters in order to have different sets of conformations for each isomer of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex. That is, a) in a first docking calculation the *Correlation Type* was changed to Shape + Electro to take into account not only the surface shape, but also the electrostatic charge distribution, whereas the *Post Processing* was changed to OPLS Minimization; b) in the second docking calculation we added the Decoy As the Reference State (DARS) [52] and we used the *Correlation Type* as Shape + Electro + DARS, while the *Post Processing* was changed to DARS Minimization; c) the third docking calculation was carried out with the parameters by default, that is, *Correlation Type* as Shape and none *Post Processing*. In all cases the number of *Solutions* was modified to 50000 and the *Final Search* parameter was changed to 30 (because the initial PDB structure had high resolution). In order to group similar structures found in the docking calculation, the structures were organized in 100 clusters, with a cluster window of 50 and a threshold root mean square (RMS) of 1.5. The rest of parameters of the clustering control panel were kept as the default. These different docking calculations were performed for each isomer, Ax and Eq, of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex and the lowest-energy systems from the docking calculations interacting though end-stacking and intercalation between DNA bases of the GQ were taken as starting points for the LS-DFT optimizations. Indeed, the obtained systems from docking were neutralized by adding K^+ close to phosphate group at a distance of 2.7 Å, and two additional K^+ were also added in the center of the GQ in the same way as suggested in the work of Housou et al. [108]. That is, each K^+ cation was coordinated

to eight O_6 from the G-tetrads of guanines, resulting in a +2 charged system. For the LS-DFT optimizations of the whole system with 966 atoms SIESTA *4.1-b3* software [7, 229] was used with the LMKLL [144] van der Waals functional, which includes dispersion corrections, being highly appropriate to characterize weak forces between the ligand and DNA. This functional was found to perform excellently for geometrical parameters and energetics of GQ [155], see Chapter 3. The system was included in a unit cell with a cell vector, in angstroms, of (60.0, 60.0, 60.0), and only the Γ k point was considered in the calculation. In order to achieve good geometries the modified Broyden algorithm was used [117], SCF was accelerated with Pulay method [188, 189], keeping the history of 4 past matrices, the density matrix mixing weight was set to 0.005, 30 meV for the energy shift and a 150 Ry mesh cut off was chosen. We also considered the SCF density matrix tolerance of $1x10^{-5}$ eV and we set the max force tolerance to 0.1 eV/Å. Double-ζ plus double polarization (DZDP) numerical basis set [119] were used in which the core electrons were substituted by norm-conserving pseudopotentials [129, 248], optimized for each element of the system with the ATOM package included in SIESTA. For a better understanding on the nature of the interaction of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex with the GQ we carried out the EDA calculations for the four most stable structures of each isomer obtained from the previous optimizations. This EDA calculations were carried out by using the ADF software [1, 87, 255]. In the EDA the interaction energy (ΔE_{int}) between fragments (DNA and metal complex) is split into different energy terms following the Morokuma energy decomposition method [126, 165] as follows:

$$\Delta E_{int} = \Delta E_{elstat} + \Delta E_{Pauli} + \Delta E_{orb}(+\Delta E_{disp}) \tag{4.1}$$

Moreover, we may define the steric contribution as [33]:

$$\Delta E_{steric} = \Delta E_{elstat} + \Delta E_{Pauli} \tag{4.2}$$

As stated by Hopffgarten and Frenking [106], if an explicit correction term for dispersion is used, the EDA results remain unchanged and the dispersion correction appears as the ΔE_{disp} extra term. In contrast, if the dispersion contribution is part of the functional, it will change the EDA results by weakening the ΔE_{Pauli} repulsive contribution. The EDA results for intercalation systems were already compared by using functionals with explicit and implicit ΔE_{disp} contributions, and similar results were obtained for all the compared final ΔE_{int} [75, 77, 78]. However, because it is more useful to have an explicit contribution for the ΔE_{disp} term in the EDA, the calculations were performed with the B3LYP-D3 functional [30, 143], which includes Grimme's dispersion corrections [84] as an extra term. We used two kinds of basis sets with different size in order to see how the EDA is affected, that is, uncontracted polarized double-ζ and triple-ζ basis sets of Slater type orbitals (DZP and TZP, respectively). For the former no frozen core approach was used, whereas for the latter a medium frozen core approximation was employed because of the size of the system (966 atoms). Relativistic effects were treated with the Zero Order Regular Approximation (ZORA) Hamiltonian [146–150]. Solvent effects were taken into account by using COSMO solvation model as implemented in ADF [127]. As far as we know, there should not be reactive processes involving the solvent and we believe that labile interactions of the hydrogen bonding produced between the solvent (water) and the structures studied in this work (GQ interacting with metal complexes) may be averaged automatically through such continuum model [247].

In order to gain insight into the nature of the interaction between the molybdenum complex including phen and the GQ, the topology of the electron density was analyzed with QTAIM [15]. Such analysis

provides an accurate definition of the chemical concept of an atom and a bond in the chemical structures by means of a partition of the molecular structure into separated atomic basins, and in consequence the concept of atom is defined inside of the molecule. The wavefunctions used for the QTAIM calculations were computed with Gaussian16 [73] at M11L/6-31+G(d,p) level of theory with the exception of the Mo atom, where we used the LANL2DZ effective core potential and the associated basis set [99] supplemented with f-polarization functions [64]. It must be said that in order to reduce the the computational cost and the generation of a huge wave function for these systems of 966 atoms, the structure of the GQ interacting with the metal complex was trimmed keeping the regions in which the Mo complex could produce BCP's with the GQ structure and removing the rest of atoms. The AIMALL software [122] was used for such QTAIM analyses. Moreover, the topology of ρ was analyzed in order to explore the non-covalent interactions by means of the NCI index, developed by Johnson et al. [118]. Such NCI analyses were also performed with the AIMALL software.

4.3 RESULTS AND DISCUSSION

4.3.1 *Geometries and energetics*

Figure 17 and 18 show the most stable optimized geometrical structures for Eq and Ax isomers of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex when interacting with the GQ through end-stacking or via intercalation between base pairs along with the relative energies, which are also depicted in Table 7 for all studied structures along with the formation energies. It must be said that whereas all most stable Ax metal complex interact via end-stacking with the GQ and the systems are quite similar with the Ax isomer localized inside the non-canonical DNA secondary structure, in the case of Eq isomers there is not any system in which the Eq isomer of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)]

metal complex is interacting via end-stacking and the conformational diversity is richer than for the Ax isomer. Indeed, for the most stable systems including the Eq isomer we found all the metal complexes interacting with the GQ from outside; two intercalating the phen ligand between the bases of the GQ, one interacting with the GQ through the Br atom and one interacting with the GQ by means of the allyl ligand.

Looking at formation energies (ΔE_{form}) and relative energies (ΔE_{rel}) (see Table 7), the most stable systems including the Ax isomer have the metal complex forming the so-called end-stacking binding mode. In this sense, more surprising is the fact that the metal complex is located in the cavity inside the non-canonical secondary structure of the DNA, completely surrounded by not only the G-tetrads of the GQ but also by the consecutive adenine tetrads. It means either that the DNA has to unfold itself to allocate the metal complex in that position and when folding back it traps the metal complex or that the metal complex promotes the folding of such non-canonical DNA secondary structures around it.

On the other hand, the most stable Ax isomers have the Br atom facing opposite to the G-tetrads therefore, moving away from the generated ion-channel, where the O_6 atoms of the guanine bases produce a high electron density concentration, which is an ideal spot to place positively charged metal ions as K^+ or other alkaline cations but not electronegative atoms such as Br. Finally, it must be said that in all cases the phen ligand is parallel to the G-tetrads, which suggests the presence of $\pi - \pi$ interactions between phen and the closest G-tetrad.

On the other hand, none of the most stable systems including the Eq isomer places it inside the DNA secondary structure. Indeed, the Eq isomer of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex is more stable either intercalating the bases of DNA from outside or interacting through the allyl or the halogen ligand with the DNA. This is an interesting result because it opens the door to the possibility of modulation of the interaction not only by substitution of the phen planar

Table 7: Obtained energy values of the formation energy (ΔE_{form}) and relative energies (ΔE_{rel}) in kcal mol^{-1} between the two isomers of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex and the GQ computed with the LMKLL/DZDP LS-DFT method.

Structure	ΔE_{rel}	ΔE_{form}	Structure	ΔE_{rel}	ΔE_{form}
Eq1	7.47	-484.22	Ax1	0.00	-491.70
Eq2	24,57	-467.13	Ax2	0.91	-490.79
Eq3	40.99	-450.71	Ax3	29.83	-461.86
Eq4	50.26	-441.44	Ax4	34.05	-457.65
Eq5	52.96	-438.74	Ax5	35.53	-456.17
Eq6	54.47	-437.23	Ax6	41.56	-450.14
Eq7	57.16	-434.54	Ax7	41.92	-449.77
Eq8	60.12	-431.57	Ax8	43.67	-448.03
Eq9	60.98	-430.72	Ax9	44.92	-446.78
Eq10	61.81	-429.89	Ax10	56.14	-435.56
Eq11	62.97	-428.73	Ax11	57.71	-433.99
Eq12	68.11	-423.59	Ax12	62.34	-429.36
Eq13	76.13	-415.57	Ax13	72.75	-418.95
Eq14	81.12	-410.58	Ax14	73.23	-418.47
Eq15	83.21	-408.48	Ax15	81.60	-410.10
Eq16	97.35	-394.35	Ax16	89.09	-402.61
Eq17	99.81	-391.88	Ax17	98.34	-393.36
Eq18	108.65	-383.05	Ax18	103.84	-387.86
Eq19	109.80	-381.90	Ax19	112.33	-379.36
Eq20	111.98	-379.72	Ax20	115.17	-376.53
Eq21	118.23	-373.47	Ax21	126.53	-365.17
Eq22	120.40	-371.30	Ax22	131.69	-360.00
Eq23	144.04	-347.66	Ax23	132.12	-359.58
Eq24	149.91	-341.78	Ax24	137.10	-354.60
Eq25	155.59	-336.11	Ax25	137.65	-354.05
Eq26	182.38	-309.32	Ax26	147.16	-344.54
Eq27	189.76	-301.94	Ax27	155.66	-336.04
			Ax28	202.91	-288.78

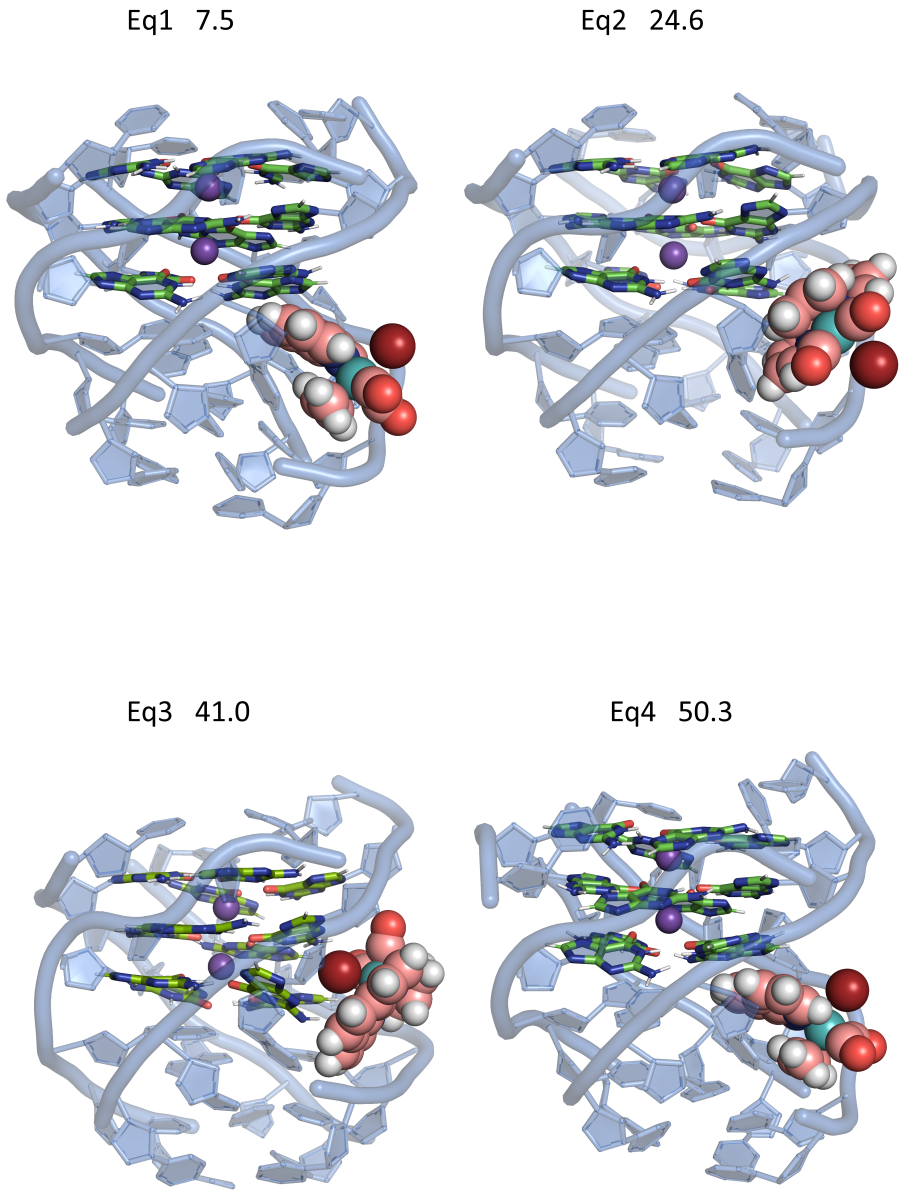

Figure 17: Most stable optimized structures along with their relative energies (kcal mol^{-1}) at LMKLL/DZDP level for the interaction between Eq metal complex and the GQ.

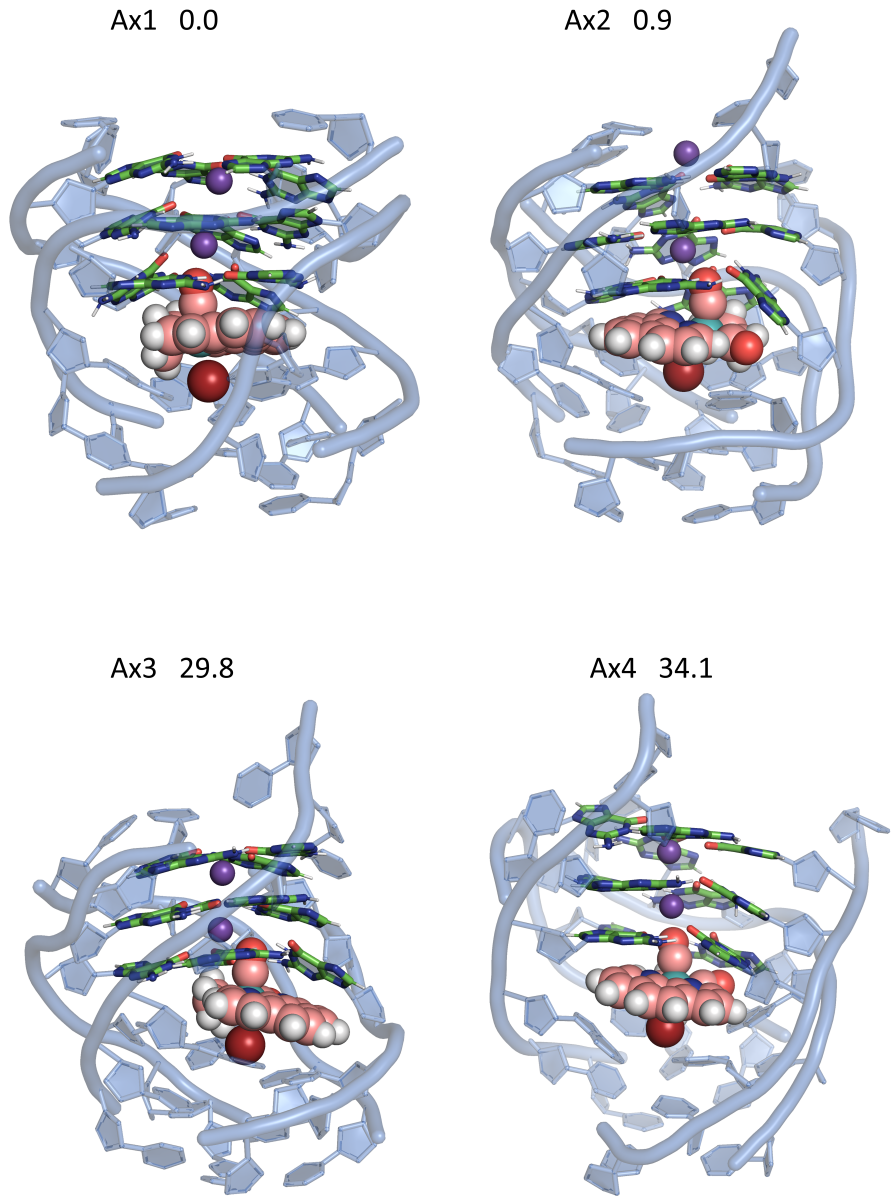

Figure 18: Most stable optimized structures along with their relative energies (kcal mol^{-1}) at LMKLL/DZDP level for the interaction between Ax metal complex and the GQ.

ligand but also by substitution of the ancillary ligands, which would have significant role in the interaction of the Eq isomer with the GQ DNA.

EDA was performed not only to determine the different contributions to the interaction energies (ΔE_{int}) between the complex and the GQ but also to rationalize the order of the stability of the structures (see Figure 19). In the case of Eq systems the most negative value was given by the Eq1 system, in agreement with relative and formation energies obtained for the same Eq system (see Figure 17 and Table 7). Nevertheless, now, the ΔE_{int} for the Ax systems results less negative than for the Eq systems as a general trend, with exception of the Eq3 system. It must be said that for Ax systems, the most negative ΔE_{int} does not correspond to the Ax1 system, which had the lowest relative energy and most negative ΔE_{form}, but to the Ax3 structure, having the Ax1 the second most negative ΔE_{int} and only differing by 1.9 kcal mol^{-1} from the ΔE_{int} of Ax3. It must be remembered that for the calculation of the ΔE_{int} the relaxation or preparation energy [106] is not considered. Taking into account that all the four lowest energy Ax structures include the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex interacting via end-stacking with the GQ inside the DNA secondary structure in a similar way, we attribute such differences mainly to the relaxation/preparation energy and to a lesser exent to the different level of calculation (LMKLL/DZDP with SIESTA vs. B3LYP-D3/TZP) with ADF).

On the other hand, all the contributions to ΔE_{int} become larger for the Ax systems than for the Eq systems with the exception of ΔE_{orb}. This is not surprising since for the Ax systems the metal complex is completely surrounded by the atoms of the non-canonical DNA secondary structure, which leads to stronger interactions. As a result, we have higher repulsive (ΔE_{Pauli}) and attractive (ΔE_{elstat} and ΔE_{disp}) contributions. At this point, it is noteworthy that the dispersion contributions (ΔE_{disp}) are important for the interaction of the Ax sys-

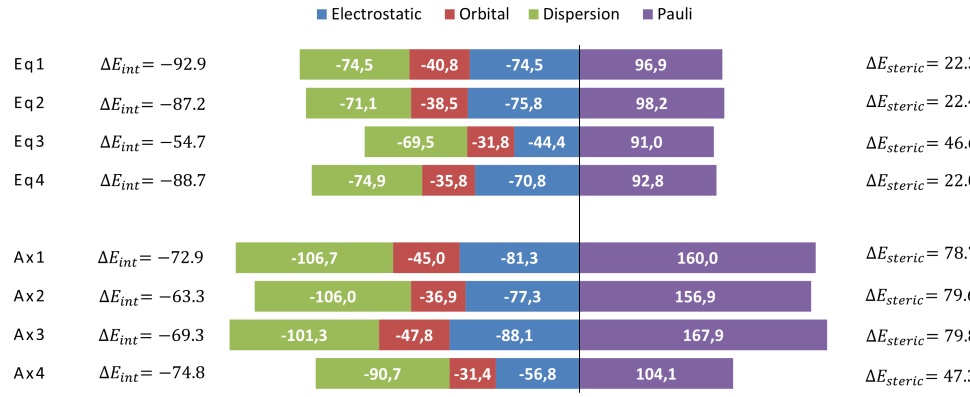

Figure 19: Cumulative bar diagram of the different contributions in the EDA at B3LYP-D3/TZP level. The total interaction energy is written in the second column and the so-called steric energy in the last column. All the energy contributions and the total interaction energy are given in kilocalories per mol (kcal mol^{-1}).

tems in the end-stacking binding mode. However, in the case of the Eq structures ΔE_{elstat} and ΔE_{disp} contributions have both a similar role in general. The ΔE_{steric} term, consisting on the sum of the repulsive ΔE_{Pauli} and the attractive ΔE_{elstat} contribution terms (eq 4.2) is very similar when comparing the systems of the same family (Eq or Ax). In the case of Ax systems, with the exception of Ax4 (47.3 kcal mol^{-1}), ΔE_{steric} ranges from 78.7 kcal mol^{-1} to 79.8 kcal mol^{-1}, with a difference of only 1.1 kcal mol^{-1}. For the Eq systems, with the exception of Eq3 (46.6 kcal mol^{-1}), the ΔE_{steric} has even lower differences among the systems and it ranges from 22.0 to 22.4 kcal mol^{-1}, which is a difference of less than 0.5 kcal mol^{-1}. Again, as stated above, we may attribute these higher values of ΔE_{steric} for the Ax systems when comparing to the Eq systems to the fact that in the Ax systems the metal complex is completely inside the cavity of the non-canonical DNA secondary structure not only interacting through the end-stacking mode with the GQ but also interacting with the subsequent adenine tetrads. In such interaction, the Ax isomer of the [Mo(η^3-

C_3H_5)Br$(CO)_2$(phen)] metal complex is surrounded by many atoms, which makes the ΔE_{steric} term higher than those obtained for the Eq, which interacts with less atoms as a general trend. The values obtained for the ΔE_{disp} are also interesting. That is, all the values obtained for the ΔE_{disp} term are more negative for the Ax isomer than for the Eq isomer (an average value of ~30 kcal mol^{-1} more negative). The more negative values of ΔE_{disp} for the Ax isomer are in agreement with the results obtained with the QTAIM and NCI analyses discussed below (Figures 20 and 21) for which more weak interactions related to dispersion forces are appreciated.

In any case, it must be highlighted that in all cases the rest of the ligands have an important role producing other weak interactions with the non-canonical DNA structure and reinforcing the interaction. In this sense, it is worth to mention that the ancillary ligands play a most important role and produce more weak interactions in the Ax systems than in the Eq structure as we will see below in the section corresponding to the analysis of the weak interactions. On the other hand, the phen ligand has the most significant differences depending if it is interacting from outside or through the end-stacking mode of interaction with the GQ inside the cavity of the non-canonical secondary DNA structure.

In order to analyze the effect of the size of the basis set, we also carried out EDA calculations with a double-ζ basis set plus polarization (DZP). Obtained energies are depicted in Table 13 of the Appendix A. It is observed that as general trend the ΔE_{int} values obtained with the DZP are more negative. That is, 30% more negative in the case of the Eq systems, whereas for the Ax systems the difference is even more considerable ($50-60$% more negative) when comparing to the results obtained with the TZP basis set. Actually, when using the DZP basis set in the EDA, the ΔE_{int} energies of the Eq and Ax systems are more similar among them, whereas when the TZP basis set is used, a

clear general trend appeared where the Eq systems had more negative ΔE_{int} energies than the Ax systems.

In addition, in the case of the Ax isomer systems, when computing the EDA's and ΔE_{int} with the DZP basis set, the system with more negative ΔE_{int} was Ax1, while for the computations with the TZP it was the Ax3 structure. In the case of the Eq systems the order was not changed and Eq1 has still the most negative ΔE_{int}. In the case of the TZP basis set, the most important attractive contribution is ΔE_{elstat} for Ax systems, whereas for Eq systems the values of ΔE_{elstat} and ΔE_{disp} are similar. In contrast, for the EDA's performed with the DZP, in all systems Eq and Ax, the most important attractive contribution is in general ΔE_{elstat}, with the exception of Ax3 and Eq3 structures, but the difference with respect ΔE_{disp} is now lower than when comparing ΔE_{disp} and ΔE_{elstat} of the Ax systems in the calculations with the TZP basis set. It is also interesting to observe how in the case of the DZP the attractive interaction ΔE_{orb} is also more negative than when the EDA is carried out with the TZP basis set (10 kcal mol^{-1} more negative in the case of the Eq systems and 20 kcal mol^{-1} more negative for the Ax structures as a general trend), which is in agreement with the trends found in the bibliography stating that the medium-size basis sets overestimate the charge-transfer energy associated to ΔE_{orb} [211, 237, 253]. Finally, the ΔE_{Pauli} repulsive contribution is more positive, as a general trend, when the EDA is performed with the DZP basis set than when it is carried out with the TZP basis set. Taking all these considerations into account, the use of a TZP basis set is justified for the correct description of the energy contributions in the EDA and the presentation of the results for this kind of systems including 1000 atoms. Even though the calculations with the TZP basis set requires more computational resources in terms of memory, disk, computing-time, etc. we were able to carry out such computations for systems with 1000 atoms in our local cluster.

4.3.2 *Solvent effects*

To gain more insight into the solvation effects in the studied process, the desolvation penalty (ΔE_{solv}) of the total interaction energy (ΔE_{int}) was calculated by means of the COSMO continuum model [127], at the B3LYP-D3/TZP level. Table 10 collects the solvation energies for the four most stable systems for each isomer of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex interacting with the GQ DNA non-canonical secondary structure addressed in this work. We observe that for such interaction the trend in the stabilization of the system may change when the solvent effects are included, as observed in previous works in which phen derivatives interacted with dDNA [65, 75]. Indeed, as was observed before in the EDA, the ΔE_{int} interaction between the studied GQ non-canonical DNA secondary structure and the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex is clearly more negative for Eq systems than in the case of the Ax structures when only the intrinsic contributions to the interaction were taken into account (ΔE_{Pauli}, ΔE_{orb}, ΔE_{elstat}, and ΔE_{disp}). However, when we consider ΔE_{aq}, defined as $\Delta E_{aq} = \Delta E_{int} + \Delta E_{Solv}$ with $\Delta E_{Solv} = E_{solv(totalsystem)} - E_{solv(GQ)} - E_{solv(metalcomplex)}$, the values of ΔE_{aq} are more negative for the systems including the Ax isomer, being Ax4 with ΔE_{aq} = -46.0 kcal mol^{-1} the most stable system. This behavior arises from the important ΔE_{Solv} penalty obtained for the systems containing Eq complexes interacting from outside with the non-canonical DNA secondary structure (from 37 to 62 kcal mol^{-1}), which are higher than the values for the systems containing the Ax isomer inside the cavity (from 27 to 35 kcal mol^{-1}). This ΔE_{solv} penalty may reverse the order of stability of the systems and thus, the inclusion of solvent effects become crucial for the final stabilization of the studied structures. Thus, consideration of solvent effects lead to more negative values for ΔE_{aq} in the systems containing the Ax isomer of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex in contrast to the most

negative ΔE_{int} energies obtained for the systems with the Eq isomer when solvation effects were not taken into account.

4.3.3 *Analysis of the weak interactions*

QTAIM topologies and the NCI index analysis provide insight into the weak interactions between the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex and the GQ. The QTAIM topologies and the values of ρ on the BCPs for the most stable studied structures are depicted in Figures 32-39 and Tables 14-21 of the Appendix B. In order to show a clearer picture only the surrounding environment of the metal complex was plotted in the figure while the remaining structure of the non-canonical DNA system was trimmed. Moreover, the NCI isosurfaces were only plotted in the areas between the metal complex and the non-canonical DNA structure, avoiding the NCI isosurfaces that appear between sugar and phosphate groups that would overload the picture. The gradient isosurfaces are colored according to the values of sign(λ_2)ρ. Isosurfaces with blue and pale green are associated to stabilizing interactions (negative values). On the other hand, yellow and red isosurfaces correspond to repulsive interactions (positive values). Very weak interactions with values close to zero are depicted in green. We show in Figure 20 and Figure 21 the NCI results for the Ax4 and Eq4 systems, respectively, which were the arrangements with more negative interaction energy for each isomer, Ax and Eq, of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex when it is interacting with the GQ DNA after the consideration of solvent effects (see ΔE_{aq} values). All the studied structures including the NCI analysis may be found in the Appendix B (Figures 40-47).

Table 8: Contributions of the solvation energies for the studied systems at the B3LYP-D3/TZP level using the COSMO approach. All the energies are given in kcal mol^{-1}. The final energy in solution (ΔE_{aq}) corresponds to the sum of ΔE_{int} and ΔE_{solv} terms ($\Delta E_{aq} = \Delta E_{int} + \Delta E_{solv}$).

System	E_{solv} (total system)	E_{solv} (metal complex)	E_{solv} (GQ)	ΔE_{solv}	ΔE_{int}	ΔE_{aq}
Eq1	-463.3	-30.1	-494.1	60.9	-92.9	-32.0
Eq2	-478.4	-29.4	-510.7	61.7	-87.2	-25.5
Eq3	-498.2	-22.8	-512.6	37.2	-54.7	-17.5
Eq4	-506.3	-29.1	-531.2	54.0	-88.7	-34.6
Ax1	-457.0	-25.1	-463.0	31.0	-72.9	-41.9
Ax2	-424.0	-22.7	-428.9	27.6	-63.3	-35.7
Ax3	-456.1	-23.5	-467.4	34.9	-69.3	-34.5
Ax4	-471.2	-22.5	-477.5	28.8	-74.8	-46.0

For the Ax4 arrangement, as we have observed above, the coordination complex is fully inserted into the DNA non-canonical secondary structure and at the same time the phen flat aromatic ligand is horizontal to the G-tetrads, whereas the ancillary ligands are confronted to the bases. This stable arrangement can be explained by means of the interactions presented in the NCI analysis. Firstly the Br atom interacts with different bases because it is located in the ion-channel of the GQ showing isosurfaces mapped in pale blue, which indicates stabilizing interactions. The same situation is found for the CO group opposite to Br that is also located in the center of the structure. This CO interacts with different atoms of the G-tetrad and also with the cation located in the center of the GQ. One of these interactions corresponds to a H atom of a G-tetrad and presents a high negative value for the $sign(\lambda_2)\rho$. Regarding the allyl group, a considerable region is shown, which corresponds to $\pi - \pi$ interactions with the nearby base plane. These results indicate that the ancillary ligands may play an important role in the overall interaction between the metal complex and the DNA. On the other hand, the surrounding area of the phen ligand presents large isosurfaces corresponding to $\pi - \pi$ stacking between the aromatic moieties. The main conclusion that arises from these results is that when the interaction of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex and the non-canonical GQ DNA structure takes place from inside, both, the Br atom and the opposite CO group interact with the polar groups of the confronted bases. This contribution to the stabilization is noteworthy due to the negative value for $sign(\lambda_2)\rho$, which shows pale blue isosurfaces. The remaining isosurfaces correspond to the surrounding areas of phen ligand and such interactions correspond mainly to the $\pi - \pi$ stacking of the phen ligand with the aromatic moieties of the G-tetrads. Moreover, the H atoms of phen show CH/n weak interactions that have been previously described in DNA ring models of phen derivatives interacting with dDNA [79, 212].

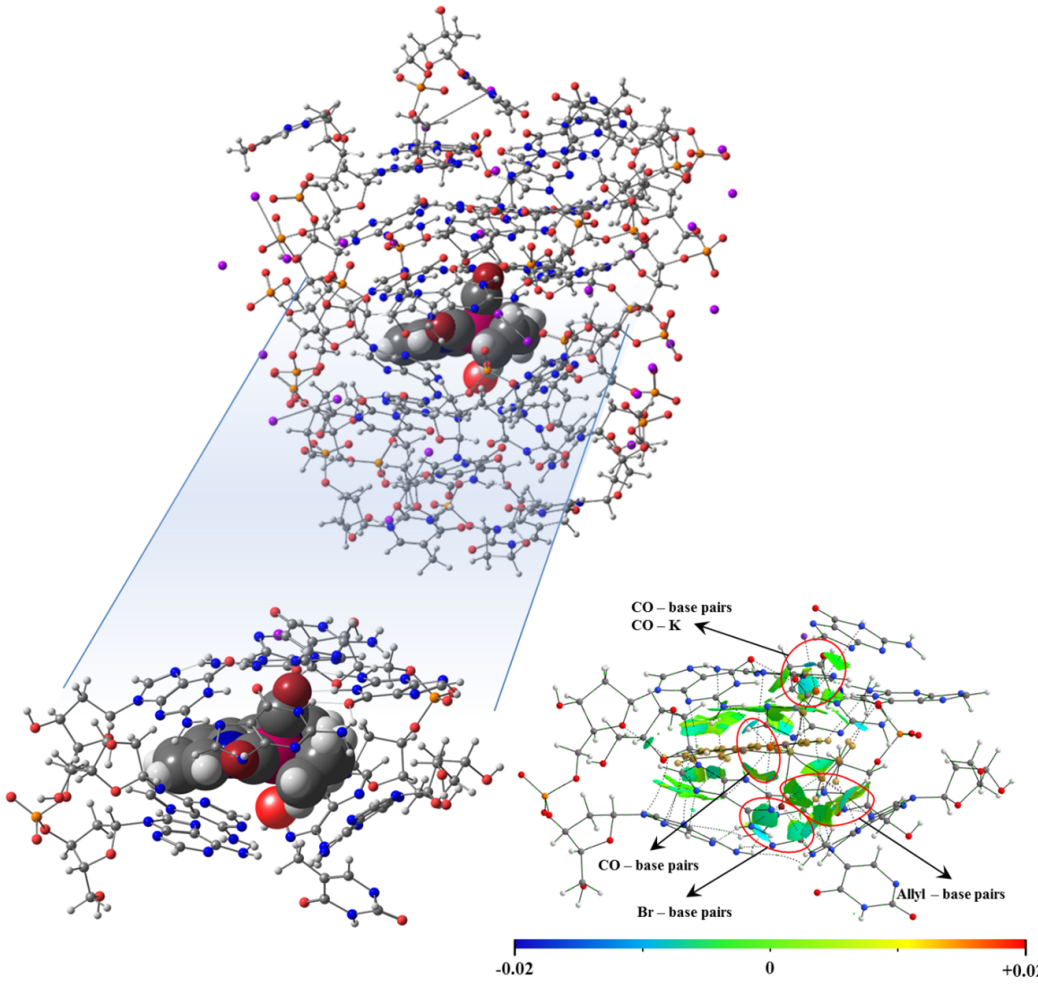

Figure 20: NCI plot with gradient isosurfaces (s = 0.5 au) computed for the considered trimmed structure of Ax4.

In the case of the Eq isomer, the most negative ΔE_{aq} energy has been found for the Eq4 system. In this geometrical arrangement the phen ligand of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex is intercalated between the tetrads, while the ancillary ligands (Br, allyl, and CO groups) remain in the outer part of the GQ, being closer to the sugar and phosphate backbone. It is observed in Figure 21 that the allyl group presents stabilizing interactions with the O atoms belonging to the phosphate groups with a considerable negative value for $(\lambda_2)\rho$ (pale blue). It must be said that the CO groups does not interact with any atom of the non-canonical GQ DNA secondary structure but with K cations that are placed in the surrounding areas of the sugar and phosphate backbone. On the other hand, it is noteworthy to say that for this arrangement the Br atom is located outside of the GQ and it does not interact with any atom of the surrounding structures being only exposed to the solvation media. Finally, regarding phen as a planar ligand, it is partially intercalated between the guanine and adenine tetrads. Such intercalation also yields the corresponding $\pi - \pi$ interactions and the CH/n interactions found previously for the Ax4.

4.3.4 *Discussion*

The main remark arising from the previous analysis is the strong influence of the location of the interaction mode, with a clear difference when intercalating between tetrads from inside or interacting from outside of the non-canonical DNA secondary structure. When the metal complex is interacting through end-stacking with the GQ totally inside the non-canonical DNA secondary structure, most of the interactions of the ligand are produced with the bases of the DNA, involving both, phen and the ancillary ligands. On the other hand, when the metal complex is interacting from outside of the non-canonical DNA secondary structures fewer $\pi - \pi$ interactions are found and the ancillary ligands interact mostly with the sugar and phosphate backbone. The results

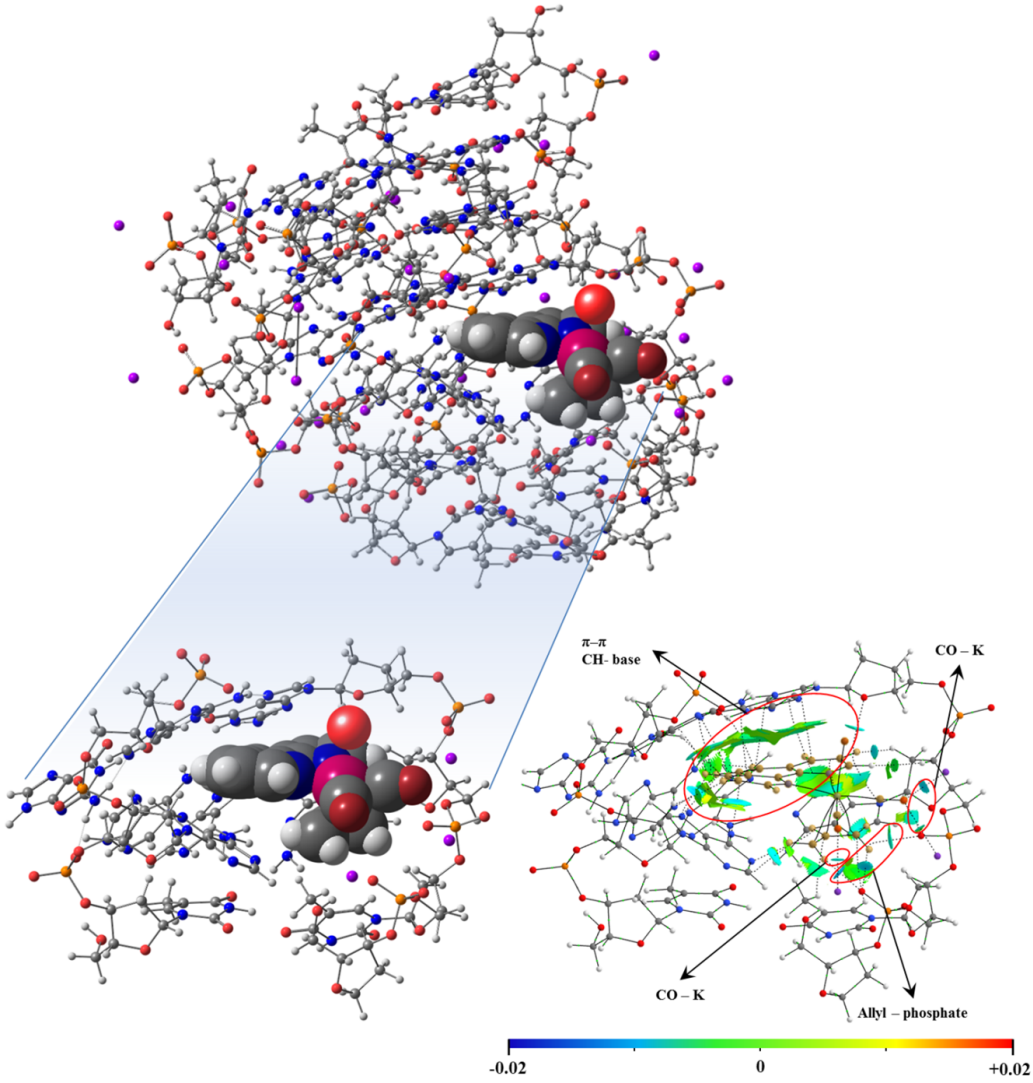

Figure 21: NCI plot with gradient isosurfaces (s = 0.5 au) computed for the considered trimmed structure of Eq4.

coming from the QTAIM and NCI analysis are in agreement with the EDA results. Our analysis may also be applied to explain the binding affinity and selectivity of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex for GQ vs. dDNA.

Regarding the formation energies of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex when interacting with the GQ, the energy difference between the binding mode achieved by each isomer, Ax and Eq, for the most stable system is not really high. On the other hand, when considering the interaction energies ΔE_{int} and ΔE_{aq} the results are very different depending whether solvent effects are considered or not. That is, ΔE_{int} is more negative when the Eq isomer of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex is interacting from outside, whereas ΔE_{aq} is more negative when the Ax isomer is interacting through end-stacking and inside of the GQ DNA secondary structure. In this sense, it should be emphasized the importance of this result, considering that in the literature no results were found in which octahedral metal complexes are interacting with the GQ inside the non-canonical secondary structure of the DNA.

The analysis based on the NCI and QTAIM shows that, end-stacking of the Ax complex with the GQ is favored due to the larger number of weak interactions than in the outer intercalation mode. In the case of Ax systems, the average value of weak interactions for the four most stable structures is 36 with an accumulated ρ value (summation of the individual values of ρ) at the BCP of 0.30 a.u. (Tables 18-21 of the Appendix B), whereas for the Eq the average of weak interactions is 33 with an accumulated ρ value of 0.25 a.u. (Tables 14-17 of the Appendix B).

The different number of weak interactions between the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex and the GQ may be also appreciated in the NCI plots (See Fig 40-47). The Ax isomer, which interacts by means of end-stacking between tetrads and it is totally surrounded by DNA atoms has an extended green surface, which represents $\pi - \pi$

stacking interactions, whereas all the ancillary ligands are also taking part in different types of interaction with the DNA. In contrast, for the Eq isomer interacting from outside of the non-canonical DNA secondary structure not all the ancillary ligands are involved in the interaction. We may summarize that the Ax isomer of the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex is more stable when interacting with the GQ through the end-staking mode of interaction due to the larger number of weak interactions, not only with phen ligand but also through the rest of the ligands. The high number of weak interactions produced by the Br ligand in its position is specially relevant because it could be easily substituted by another halogen atom or even another group like triflate (Tf) [20], which would produce the modulation of the interaction with the GQ DNA.

4.4 CONCLUSIONS

In this work, for the first time, we have studied different interaction modes for different isomers, namely, Eq and the Ax, of the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex when interacting with non-canonical GQ DNA secondary structure, containing 966 atoms by means of a LS-DFT method. We mainly analyzed two interaction modes, the end-stacking mode, where the complex is completely inside the non-canonical DNA structure, and the interaction from outside with DNA bases in which intercalation of the phen ligand is the main mode of interaction. The Ax isomer prefers the end-stacking mode of interaction, while for the Eq isomer the intercalation from outside the non-canonical DNA secondary structure is favored. The combined results from geometry optimizations, formation energies, EDA including solvent effects, QTAIM, and NCI analyses lead to the conclusion that the end-stacking mode of interaction of the Ax isomer between tetrads of DNA is the most stable binding mode. However, one should take into account that to access the intercalation binding mode the GQ

should be unfolded for the latter introduction of the metal complex inside the non-canonical DNA structure. Once unfolded, the presence of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex could act as promoter and stabilizer for the formation of such GQ.

Subsequent analyses by means of EDAs and introduction of solvation energies confirm that although both interaction modes are stable, solvent effects favour end-stacking of the Ax isomer. Moreover, QTAIM and NCI analyses confirm that this interaction of the Ax complex produces a larger number of weak interactions than the outside intercalation binding mode of the Eq isomer.

5

BINDING PREFERENCE OF MOLYBDENUM (II) PHENANTHROLINE COMPLEXES BETWEEN DUPLEX AND G-QUADRUPLEX DNA

Up to now, we have discussed the performance of different methods when describing GQ (Chapter 3) and their interaction with two isomers of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex (Chapter 4). In this chapter, we will discuss the binding affinity and selectivity of the Mo-based metal complex between the dDNA and GQ.

5.1 INTRODUCTION

One of the main problems when fighting against tumoral cells is the selectivity of the drugs. The drugs that we use against the disease do not distinguish between tumoral and somatic cells, indiscriminately attacking any DNA molecule, which causes serious side effects and damage. The interest on targeting GQ structures has increased, since their relationship with cancer has been revealed. GQ's have been associated to a large number of genomic functions, such as translation and replication [111] and the studies have revealed their presence in key regions of the human genome, such as promoters [224] and telomers [34]. In the case of telomers, the formation of telomeric GQ has been shown to inhibit telomerase activity [71] which is responsible for maintaining the length of the telomers and it is involved in the 85% of cancers [159]. This inhibition causes apoptosis of cancer cells and, since telomerase is overexpressed in the majority of cancer cells and in relatively few somatic cells, it is recognized as a potential cancer-specific target that would avoid the death of somatic cells in chemotherapy treatments for

tumors. For this reason, this strategy has been considered in current drug design to devise alternative therapies for cancer treatments [154]. In this sense, the studies of small molecules that bind and stabilize GQ have recently become a hot topic of pharmaceutical and medical research. Ideally, one aims to synthesize and produce different molecules that have not only affinity but also selectivity towards GQ over dDNA, and at the same time, they are capable of stabilizing GQ's. One of the drawbacks when facing this problem is that interacting molecules must be able to distinguish between different structures of the same DNA, with similar binding modes in both cases. That is, structures present intercalation and groove binding interaction modes through the mg or the MG, while for the GQ side loop and end-stacking binding modes are also present. There are also some differences between these binding modes that could be exploited in order to create selectivity. The main difference between both DNA structures is that dDNA is made up of two helix-shaped strands, while GQ are made up of four strands (they do not have to come from different chains). In addition, the core of the GQ is formed by the stacking of G-tetrads, composed of four guanine bases interacting through Hoogsteen H-bonds, while dDNA is always composed of base pairs, interacting mainly through the Watson-Crick H-bonds.

As mentioned before, there are some characteristics for ligands that increase the affinity for GQ over the dDNA: (*a*) a π-delocalized system with a surface area similar to the G-tetrad; (*b*) a positive partial charge, which interacts with the O6 of the guanines oriented towards the center of the one in G-tetrads; and (*c*) positively charged subsituent arms, which can interact with the negatively charged phosphate backbone. Taking this into account, most of the studied ligands are organic planar molecules with high number of aromatic rings. However, metal complexes, although less studied than organic ligands, offer a wide range of structural and electronic properties that can be used when designing new drugs.

Among the organic molecules, telomestatin [125], hexaoxazoles [23], and bistriazoles [161] have been shown to have high affinity for the GQ and exhibit moderate levels of cytotoxic activity. Pophyrins, which have a wide aromatic surface, and their derivatives have also been extensively studied. Metalloporphyrins were one of the earliest synthesized metal complexes, due to the assumption that the introduction of a metal ion would add some electrostatic effects that would enhance binding affinity. One of the most efficient pophyrin-based ligand is Mn^{III}-porphyrin, which can distinguish bethween dDNA and GQ by four orders of magnitude [62]. Planar metal complexes made of combinations of Ni^{II}, Cu^{II}, and Pt^{II} metal atoms and both salphen and salen ligands are one of the most studied complexes, which has been reported as efficient GQ binders, being able to intercalate between the DNA bases through $\pi-\pi$ stacking [121, 241, 261]. Ru^{II}, Ir^{III}, and Fe^{III} octahedral complexes with planar ligands, such as ethylenediamine, bipyridine, phenanthroline, dipyridophenazine, phenanthroimidazol, and phenylpyridine, have also been investigated, but especially ruthenium polypyridil complexes. In most cases, the planar ligands with large π-delocalized aromatic surfaces are involved in end-stacking or intercalation between DNA bases. However, it is more difficult for the non-planar part of the molecule to interact through these binding modes due to geometric limits, leading into a problem for the achievement of an efficient interaction.

In the previous Chapter 4, the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ metal complex has demonstrated to be a promising GQ binder. However, it is possible that this metal complex would show a poor selectivity for the GQ compared to the dDNA, even that it show a good affinity towards GQ. For this reason it is important to compare and rationalize the binding affinity and selectivity of the metal complex for both secondary DNA structures. In this chapter we will discuss the binding selectivity of the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ metal complex between the dDNA and

GQ, defined as the difference between the corresponding formation interaction energies of the metal complex with the GQ and dDNA.

5.2 METHODS

The methodology used to describe the interaction of the metal complex interaction with the GQ has been described in chapter 4. The following subsection corresponds to the methodology used to study the interaction of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex with dDNA. The model to study the interaction of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex with the DNA octamer was built from the 1N37 structure of the PDB, reported by Searle et al. [218]. It consists of the respinomyocin D drug, from the family of anthracycline antitumor antibiotics, which interacts by intercalation with the d(AGACGTCT)$_2$ DNA canonical sequence between CG/GC base pairs. The original drug, respinomycin D, was removed, creating a pocket to host the Mo complex. Because different kinds of interactions will be studied and they may appear in different regions of the DNA d(AGACGTCT)$_2$ sequence, a strategy was developed to find the lowest-energy systems, using docking to generate starting points for the LS-DFT optimizations. For the intercalated systems, docking calculations were carried out by taking the d(AGACGTCT)$_2$ DNA sequence of the 1N37 PDB structure, where respinomyocin D was substituted by Ax and Eq [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex, considering both isomers. The HEX *8.0.0* software was used [166] following the HEX manual recommendations. Due to the the characteristics of our system (DNA chain + metal complex), we took the default parameters of the docking control panel for the docking calculations and we modified only (1) the *Correlation Type* (changed to Shape + Electro to take into account not only the surface shape but also the electrostatic charge distribution), (2) the *Post Processing* (OPLS Minimisation instead of the no Post Processing

default), (3) the default number of *Solutions* (50000 solutions), and (4) the *Final Search* (30 because the PDB structure had a high resolution).

The structures from molecular docking were neutralized by adding as many Na^+ cations as the number of phosphate groups. These Na^+ cations were put 2.8 Å far away from the phosphate O atoms, and LS-DFT geometry optimizations were carried out for the most representative structures of the docking calculations by means of SIESTA [229]. These optimizations were performed for the whole system of 556 atoms. The system was included in a unit cell with a cell vector, in angstroms, of (34.17, 31.36, 32.03), and only the Γ k point was considered in the calculations. Given the importance of weak interactions in these kinds of biological systems, as in the previous chapter, we used the LMKLL functional, which includes van der Waals contributions [144], with the double-ζ double-polarization (DZDP) numerical basis sets [119] with core electrons were substituted by norm-conserving pseudopotentials [129, 248]. The cutoff radii for the atomic orbitals of each element were obtained for an energy shift of 30 meV. The tolerances used in the optimizations were 10^{-5} eV for the energy and 0.02 eV/Å for the forces. The EDA was performed with the B3LYP-D3 functional with Grimme's D3 correction to dispersion forces [30, 84, 143] and an uncontracted polarized triple-ζ (TZP) basis set of Slater-type orbitals. Solvent effects (water) were taken into account by using the continuous solvent model COSMO [127], as implemented in the ADF software [1, 87, 255].

In order to achieve a proper description of the most stable structures for each interaction mode of $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ metal complex with dDNA, a QTAIM [15] topological analysis of the electron density, ρ, was carried out to identify BCP's and bond paths with the AIMALL software [122]. The wavefunctions for that were computed with the Gaussian16 software [73] at the M11L/6-31+G(d,p) level of theory for all the atoms except for Mo, for which we used the Hay-Wadt LANL2DZ [99] effective core potential and the associated basis

set, to which a set of f polarization functions was added [64]. AIMALL software was also used to calculate the NCI index [118], based on the peaks that appear on the reduced density gradient mapped versus the electron density multiplied by the sign of the second Hessian eigenvalue, $sign(\lambda_2)\rho$, which complements the study of the weak interactions. A reduced model of the structures containing the metal complex and DNA was considered in the QTAIM and NCI analyses to save computation time. Thus, the $d(ACGT)_2$ tetramer was used for the intercalation model.

5.3 BINDING SELECTIVITY DIFFERENCES BETWEEN DUPLEX AND G-QUADRUPLEX DNA FOR THE INTERACTION WITH THE MO-BASED METAL COMPLEX

5.3.1 *Energetic Results*

In previous works, $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ metal complex demonstrated to have a good binding affinity with dDNA [65]. Optimized geometries and formation energies obtained at LMKLL/DZDP level of theory (Figure 22) demonstrated that the intercalation via mg is the most favourable mode of interaction for both metal complex isomers, Eq and Ax, being at least 12.7 and 15.9 kcal mol^{-1} more favourable than the intercalation though the MG, respectively. Moreover, most of the obtained formation energies for the Ax isomer were slightly more negative than their Eq counterparts. On the other hand, the mg binding mode of interaction (without intercalation) was the less favourable one, with the less negative formation energy. Both isomers of the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ metal complex have a clear binding affinity for the DNA structures, since the LS-DFT formation energies and EDA interaction energies are negative (see figures 17 and 18 for the GQ, and 22 for the dDNA). In the case of the interaction with the dDNA the formation energies for the most stable Ax isomer systems

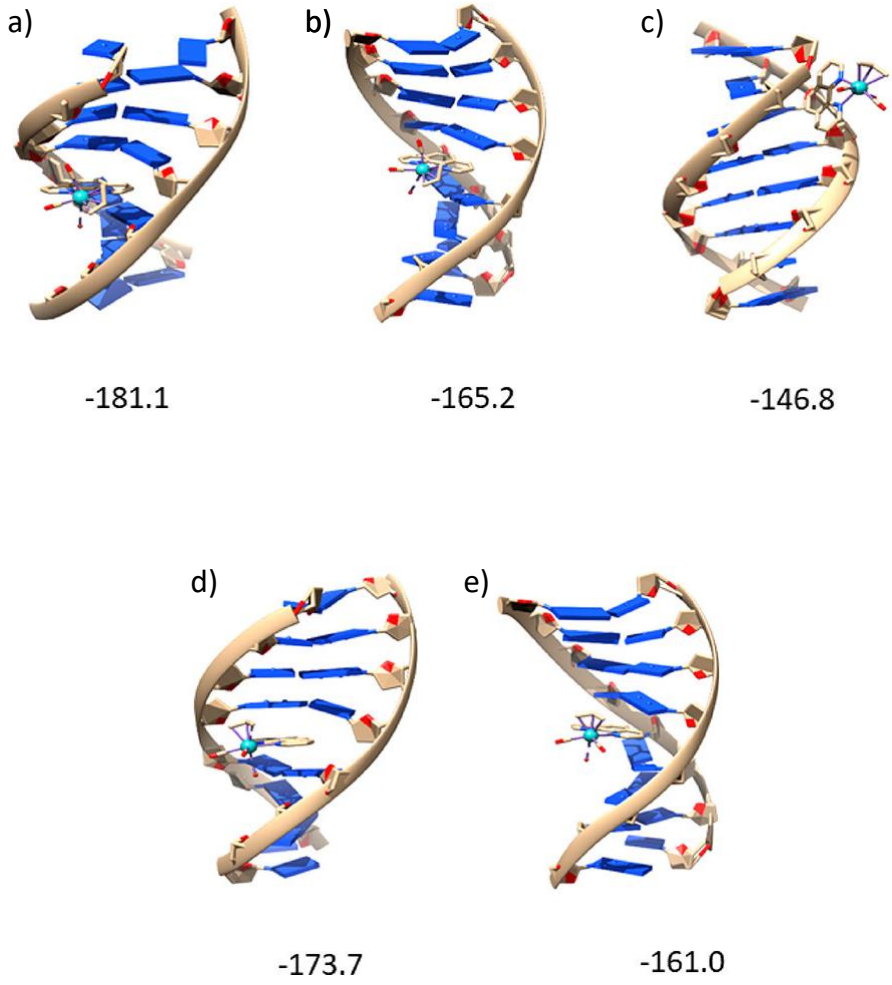

Figure 22: Most stable optimized structures along with their formation energies (kcal mol^{-1}) at LMKLL/DZDP level for the Ax metal complex a) intercalated with dDNA through the mg, b) intercalted with dDNA through the MG, c) interacting with dDNA thught the mg binding; and for the Eq metal complex d) intercalated with dDNA through the mg and e) intercalated with dDNA through the MG.

they range between -181.1 and -146.8 kcal mol^{-1}, and for the Eq isomer systems range between -173.7 and -161.0 kcal mol^{-1}, while for the GQ formation energies range from -484.2 to -429.9 kcal mol^{-1} for the Eq isomer of the metal complex and from -491.7 to -448.0 kcal mol^{-1} for the Ax isomer. All these negative fomation energies obtained for the metal complexes interacting with the two secondary DNA structures lead us to the conclusion that the metal complexes have good affinity with the two DNA structures. However, based on the formation energies between dDNA and GQ, on average -160 and -460 kcal mol^{-1}, respectively, one can conclude that the metal complexes have a selectivity for the GQ DNA secondary structure.

In addition, EDA calculations were carried out to determine the ΔE_{int} interaction energy and the different contributions of it (Table 9). A similar trend between ΔE_{int} energies and fomation energies were observed, in that ΔE_{int} was always more negative when intercalation takes place through the mg. Moreover, it was concluded that the dispersion energy term, ΔE_{disp}, plays a crucial role in this preference. The ΔE_{disp} attractive contribution is comparable to the destabilizing ΔE_{Pauli} contribution in most cases. The more negative values of ΔE_{disp} and ΔE_{int} when comparing to isolated phen ligand and and derivatives interacting with dDNA [77–79] can be associated to the additional interactions of the other ligands of the metal complex with dDNA, which involves, Br, allyl, and CO, which are not spectator ligands but play a significant role in the interaction of the complex in the intercalation pocket with dDNA by means of weak interactions. These weak interactions are not only with the base pairs but also involve the sugar and phosphate backbone.

It is observed that the ΔE_{int} with EDA, are slightly more favourable for GQ systems than those obtained for dDNA, which combined with the information obtained from the LS-DFT formation energies point to a certain selectivity of the metal complex with respect to the GQ structures over the dDNA. The ΔE_{int} values for GQ systems are between

Table 9: Different contributions in the EDA at the B3LYP-D3/TZP level for the Eq and Ax isomer interaction with the dDNA. Energy contributions are given in kcal mol^{-1}

System	ΔE_{int}	ΔE_{disp}	ΔE_{orb}	ΔE_{elstat}	ΔE_{Pauli}
Ax Intercalation minor groove	-57.6	-59.1	-19.5	-36.7	57.7
Ax Intercalation major groove	-50.5	-53.1	-19.5	-34.8	57.0
Ax minor groove binding	-40.9	-36.3	-14.5	-27.1	37.0
Eq Intercalation minor groove	-55.1	-59.6	-21.6	-38.4	64.5
Eq Intercalation major groove	-50.1	-43.6	-19.2	-35.8	48.4

-54.7 and -92.9 kcal mol^{-1}, while the values for the dDNA range between -40.9 to -57.6 kcal mol^{-1}. Regarding the repulsive enery term, ΔE_{Pauli}, it is observed that for the intercalation of the metal complex it ranges between 48.4 and 64.5 kcal mol^{-1}, whereas for the GQ systems and the most stable Ax metal complex, they range between 104.1 and 167.9 kcal mol^{-1}. Thus, the ΔE_{Pauli} is 2/3 times more repulsive in GQ systems. The reason for this behavior is that when the metal complex is accommodated within the GQ it is surrounded by a larger number of atoms than in the case of dDNA intercalation, where only the phen ligand of the metal complex is completely surrounded by atoms of the dDNA. In the case of the Eq complex interacting with the GQ, the repulsion is not so large. Being intercalated by one side only phen is completely surrounded by atoms, similar to the structure in dDNA. However and unlike dDNA, the distance between sugar chains is much shorter, and therefore the binding pocket is smaller, resulting in a larger repulsion than in the case of the dDNA intercalation. For the rest of the attractive terms, namely, ΔE_{disp}, ΔE_{orb}, and ΔE_{elstat}, a similar behavior occurs to the one observed for the ΔE_{Pauli}. Namely, the Ax systems, completely inserted within the GQ, show a larger interaction with both the surrounding guanine bases and with the sugar and phosphate backbone. Due to that, the attractive contributions are more than twice larger for interactions with GQ than for interaction with dDNA. Among the attractive contributions, the dispersion term

is the one that presents the highest differences. The Ax systems, completely inside the GQ, present a significantly more stabilizant ΔE_{disp} values, whereas the rest of attractive contributions are similar in both cases.

Solvent effects were also taken into account when studying the interaction between the metal complex and the dDNA. The resultant ΔE_{aq} energy values are more similar between the different systems, although the Ax isomer intercalation via mg is still the most stable structure. The consideration of solvent effects further support a higher affinity for GQ structures of the metal complex. Thus, the ΔE_{aq} for GQ structure range between -46.0 and -17.7 kcal mol^{-1}, and between -24.0 and -5.5 kcal mol^{-1} for dDNA. That is, ΔE_{aq} are at least -10 kcal mol^{-1} more favourable for GQ than for dDNA.

Table 10: Contributions of the solvation energies for the studied dDNA systems at the B3LYP-D3/TZP level using the COSMO approach. All the energies are given in kcal mol^{-1}. The final energy in solution (ΔE_{aq}) corresponds to the sum of ΔE_{int} and ΔE_{solv} terms ($\Delta E_{aq} = \Delta E_{int} + \Delta E_{solv}$).

System	E_{solv} (total system)	E_{solv} (metal complex)	E_{solv} (duplex DNA)	ΔE_{solv}	ΔE_{int}	ΔE_{aq}
intercalation via the mg Ax	-473.6	-25.8	-481.4	33.6	-57.6	-24.0
intercalation via the mg Eq	-476.2	-27.3	-481.6	32.7	-55.1	-22.4
intercalation via the MG Ax	-480.9	-26.2	-481.9	27.2	-50.4	-23.2
intercalation via the MG Eq	-476.9	-27.4	-481.1	31.6	-50.2	-18.6
mg bindinng Ax	-457.9	-24.4	-468.9	35.4	-40.9	-5.5

5.3.2 *Analysis of the weak interactions*

To understand the nature of the interaction and the affinity between the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex and both secondary DNA structures (GQ and dDNA) we carried out QTAIM and the NCI index analysis (Figure 23). The total number of interactions obtained by QTAIM in all cases is higher for GQ systems than for dDNA, having 34 and 24 weak iteractions on average, respectively. Besides, in the NCI analysis, one can observe not only a higher number of isosurfaces, but also larger ones in the case of GQ. In addition, for the dDNA systems, in the same way as for the GQ, it is observed that the Ax isomer (figures 44-47 in the Appendix B) presents a larger number of weak interactions than the Eq isomer (figures 40-43 in the Appendix B). Analyzing the results in more detail, phen presents for both secondary DNA systems, GQ and dDNA, a large flat isosurface between phen and bases, corresponding to the $\pi - \pi$ interactions, but the isosurfaces are slightly larger for the GQ systems. This is because the Ax isomer is completely inside the cavity of the GQ and phen can create weak interactions with four bases of DNA, while in dDNA only does it with one base pair.

Equally important is the contribution of ancillary ligands to the discussion of the affinity of the metal complex with DNA. When analyzing the weak interactions formed by the ancilliary ligands, the first result that can be observed is that among all the studied systems, the highest number of interactions are formed between the Ax isomer and the GQ. The QTAIM shows a very high number of weak interactions in which the ancillary ligands are involved, and it is confirmed by the NCI where multiple isosurfaces can be observed around all ancillary ligands. All ancillary ligands (Br, CO, and the allyl group) have a high activity with DNA bases, not only with the bases that form the closest tetrad but also with those in the subsequent tetrad. It is very interesting to highlight that in all the cases it is observed that the CO ligand presents some interaction with the K^+ of the ion-channel. It

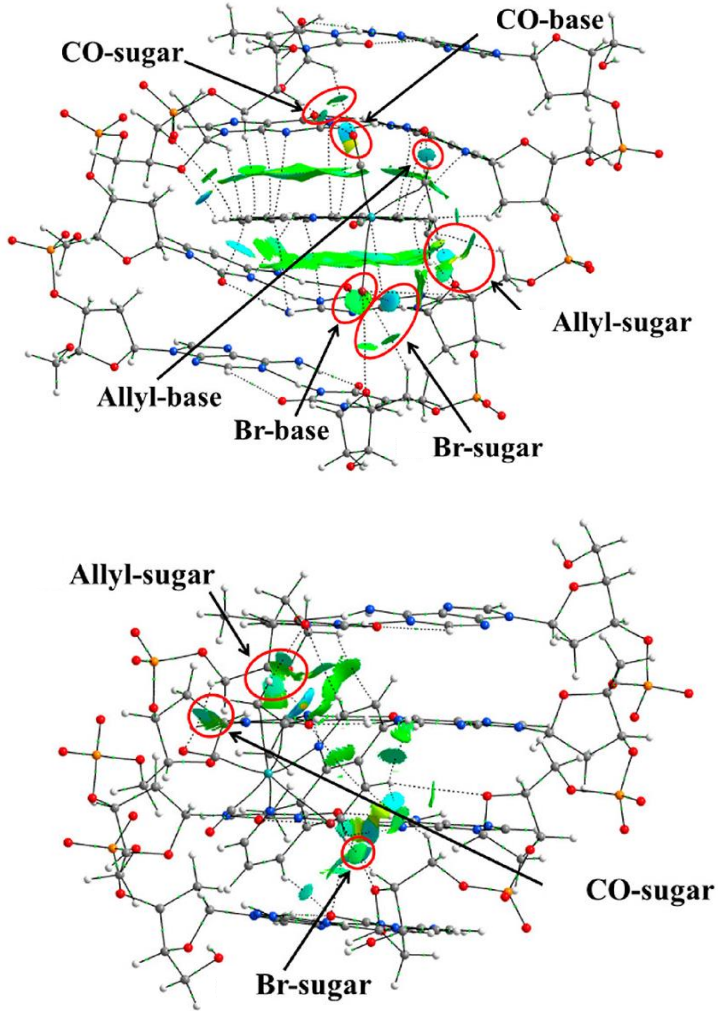

Figure 23: NCI index plots with gradient isosurfaces (s=0.5 au) computed for the considered DNA models of the most stable Ax isomer intercalated via mg (top) and Ax isomer interacting via mg binding mode (bottom) structures.

is also worth to mention that the Ax isomer is completely inside and centered between the DNA tetrads, all the formed weak interactions are between the metal complex and DNA bases, except for few interactions between the allyl group and the sugar and phosphate backbone.

The weak interactions observed between the ancillary ligands of the Eq isomer and the GQ differ from those observed for the Ax isomer due to the fact that the metal complex in this case is not found inside the DNA cavity. Correspondly, fewer weak interactions are stablished with the DNA bases and even some ligands do not create any, but high number of weak interactions are observed with the sugar and phosphate backbond.

In the case of the metal complex interacting with dDNA, as the interaction is always throught one DNA side, not only fewer number of interactions are observed between the ancillary ligands and the DNA, but also in some cases one ancillary ligand does not even interact with the DNA.

In summary, we can say that the metal complex will have a higher selectivity for the GQ, because of the higher number of weak interactions, but also because it presents more promising attractive weak interactions.

5.4 CONCLUSIONS

In this chapter we have compared the affinity and selectivity of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex between dDNA and GQ. In both cases, the obtained formation energies by means of LS-DFT as well as the interaction energies obtained with EDA point to a stabilization of the stability of the system when the interaction between the metal complex and DNA is produced. Thus, both DNA structures, dDNA and the GQ have good affinity for [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)]. Moreover, the QTAIM and the NCI analyses confirms the same type of weak interactions for both dDNA and GQ, which show $\pi - \pi$ stacking interaction be-

tween the phen and the DNA bases, whereas the ancillary ligands play also a key role in the final stabilization. Taking into account all these observations, we can affirm that the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] complex may interact with both DNA secondary structures, the dDNA and theGQ. However, LS-DFT formation energies, EDA interaction energies and the characterization of the weak interactions by QTAIM and NCI, give a coherent picture in which the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] interact more strongly with GQ structures than with dDNA, suggesting a good selectivity for GQ structures. This opens the possibility to use this type of metal complexes as promising candidates for more selective drugs that bind to the GQ and disrupt the telomerase activity.

The stabilization is higher in the case of GQ and therefore this [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex could be selective for this non-canonical DNA secondary structures.

6

SUMMARY AND FINAL REMARKS

G-quadruplexes have gained interest in the last decades due to its possible use to replace current antitumor therapies, which present harmful side effects to the human body. Additionally, they could also be used to combat antibacterial resistance. Studying these secondary structures is key to understanding the biological mechanisms involved and how they affect them effectively. Investigations of GQ fall into two main categories: i) studies of the behavior of the DNA chain when it adopts the secondary DNA structure of GQ, as well as the process of folding and unfolding of the structure, and ii) studies of the interaction of GQ structures with small molecules, purely organic molecules or metal complexes. In this thesis, we have focused on the second type of study. When studying the interaction of GQ with small molecules, the most common goal is to stabilize the GQ structure as much as possible. This stabilization disrupts different biological processes, such as transcription or replication, leading to cell apoptosis, affecting specially tumoral cells due to their fast replication. There is clear biophysical evidence of small molecules binding to the GQ, with variable selectivity for the dDNA or GQ sequences. Although the effects of such small molecules on specific cellular mechanisms (growth inhibitory effects, disruption of the transcription and replication, or alteration of telomerase activity...) are promising in terms of their development into potential anticancer therapeutics, several future challenges must be addressed to confirm the robustness of the GQ targeting concept. Studying these secondary DNA structures can benefit from computational studies, which help to understand the nature of the ligands' interaction. Computational methods have experienced remarkable growth in

recent decades, with machines with more computing power and more efficient algorithms leading to more accurate and faster techniques.

Until recently, classical molecular mechanics and quantum mechanics methods were the two main computational techniques to study biological systems. The first can easily handle large biological systems of thousands of atoms for simulations lasting from picoseconds to several milliseconds. However, they do not explicitly describe the electronic structure of the molecules, relying on highly empirical energy functions that may not be optimum to reproduce non-canonical biological structures, such as GQ. On the other hand, quantum methods are highly accurate computational methods that explicitly describe the electronic structure. Still, they are computationally expensive, and therefore, they can only describe systems of few atoms, and the simulations are only performed for short time scales. In the last decade, we have been closing the gap between these methods thanks to linear-scaling methods, which make it possible to perform calculations for large systems, such as most biological systems, without sacrificing accuracy and describing the electronic structure of the entire system. A linear-scaling method is nothing more than a computational method in which the computation time increases linearly with the size of the system (O(N) scaling). Using these methods in conjunction with DFT and CCSD allows for accurately modeling large-scale systems with a reasonable computational time. One of the objectives of this thesis was to demonstrate the accuracy with which different linear-scaling methods could characterize various large-scale secondary DNA systems and their interaction with small molecules. The different DNA structures used in this thesis have been: 1) DNA base pairing, interacting through hydrogen- bonds and stacking; 2) models of dDNA interacting with a phen ligand; 3) a complete dDNA system interacting with different small molecules, Respinomycin D and 2 isomers of the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ metal complex; 4) various structures of two G-tetrad models of GQ systems, stabilized with various alkaline

metals, and 5) A complete GQ system interacting with different ligands, such as the MMQ-1 quinacridine or 2 isomers of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex. We compared the performance of two SE methods incorporating dispersion corrections, PM6-DH2 and PM7, QM/MM method at B3LYP-D3(GD3BJ)/6-31+G(d,p):AMBER and M11L/6-31+G(d,p):AMBER levels of theory, DLPNO-CCSD(T)/def2-SVP, and LS-DFT implemented in SIESTA and using the LMKLL/DZDP level of theory.

We have shown that the PM6-DH2 SE method correctly describes the geometries of these systems. On the contrary, the PM7 Hamiltonian gives worse results and does not accurately reproduce the GQ structures. However, the two semi-empirical methods have difficulties describing these systems' interaction energies, giving contradictory results to those obtained by the more accurate methods. On the other hand, we have shown how the LMKLL functional applied through SIESTA (LS-DFT) reliably predicts both the geometry and the interaction energies for these DNA systems at a reasonable computing time. Furthermore, all the obtained energies were comparable to those obtained with the novel near linear- scaling method DLPNO-CCSD(T) carried out using ORCA package. This method can perform single-point calculations for hundreds of atoms at the CCSD(T) level of theory, and their results can be considered benchmark data. Our work demonstrates the adequacy of LS-DFT methods for computational studies of large DNA systems. This is particularly interesting to describe the interaction between small ligands and DNA, since an accurate description of the different non-covalent weak interactions, polarization, and charge transfer is needed to characterize the system and its interactions with small molecules. It also opens the field of possible development of algorithms that, through the application of linear-scaling techniques, are capable of obtaining accurate energetic results at the CCSD(T) level on structures optimized previously with LS-DFT.

Once we demonstrated that the methods are able to describe dDNA and GQ systems and their interaction with small molecules accurately, we proceeded to study the interaction between the GQ and two isomers of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex. Our study focuses on two interaction modes: the end-stacking binding mode, where the complex is entirely inside the GQ, and the intercalation binding mode. For the energetics, we used the LS-DFT methods to calculate the formation energies and the EDA to determine the role of the different energetic contributions in the interaction energy and the solvent effects. We concluded that the Ax isomer prefers the end-stacking mode of interaction, while Eq isomer favors the intercalation. In addition, we performed QTAIM and NCI studies with the AIMALL software, which provide insight into the weak interactions between the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex and DNA. Based on the NCI and QTAIM analysis, end-stacking of the Ax isomer should be favored owing to a larger number of weak interactions. The Ax isomer, interacting through the end-stacking between tetrads and surrounded by DNA atoms, exhibits many $\pi - \pi$ interactions. At the same time, all the ancillary ligands of the metal complex also establish relevant interactions with DNA. In contrast, in the case of Eq isomer interacting from outside the non-canonical DNA secondary structure, not all the ancillary ligands participate in the interaction.

All these results support a substantial affinity of the metal complex with GQ. To complete our study, we also wanted to determine the selectivity of the metal complex towards dDNA or GQ. In other words, the relative affinity between dDNA and GQ for this type of metal complexes. To do so, we compare formation, interaction, and interaction energies taking into account solvation energies, and analyze weak interactions with both DNA secondary structures. The ΔE_{int} results obtained with the EDA reveal a more favorable interaction between the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex and the GQ DNA than with the dDNA. The most stable interaction among all the systems is

between the Ax isomer and GQ DNA. The Eq isomer also shows more negative interaction energy with GQ DNA than with the dDNA. Still, the energetic gap between the two secondary DNA systems is not as big as in the case of Ax isomer systems. The results with solvation effects reinforce the idea that the metal complex would have more selectivity with the GQ DNA since the ΔE_{aq} are 2/3 times higher than for dDNA. Moreover, the analysis of the weak interactions, using QTAIM and NCI, adds more evidence to the hypothesis that the interaction of the metal complex would be more favorable with the GQ DNA. The analysis of the weak interactions reveals that the interaction of the Ax isomer with the GQ DNA, entirely inside the cavity of GQ and surrounded by DNA atoms, creates the widest $\pi - \pi$ interaction, taking maximum advantage of the surface between the phen and the tetrads. Additionally, ancillary ligands are also involved, creating relevant interactions with the tetrads, not only with the closest ones but also with those of the subsequent tetrad layer. Furthermore, Eq isomer interacting from outside the GQ DNA has also shown large $\pi - \pi$ surface areas. Although its ancillary ligands are not as active as in the case of Ax isomer, they have a crucial role in the interaction. On the other hand, the interaction of the metal complex with the dDNA is not as favorable as for the GQ DNA. It was observed that for the Ax isomer there is a wide $\pi - \pi$ surface, but only between base pairs, and therefore, it is not as wide as in the case of GQ DNA. In addition, since only the phen ligand intercalates between DNA bases and the remaining ligands lie outside, not all the ligands are involved in the interaction. The weak interactions observed for the Eq isomer do not promote the selectivity for the dDNA. The $\pi - \pi$ isosurface was not very wide because the metal complex was not completely intercalated. Besides, some of the ligands did not take part in the interaction, resulting in the least promising interaction mode of all those analyzed. Taking all into account, we conclude that the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex will have an affinity with the two DNA secondary structures and will stabilize both of them.

Still, it will have selectivity with the non-canonical GQ DNA secondary structure since it has a more significant affinity for GQ DNA than for dDNA.

7

SARRERA

Kantzerraren kontrako terapiak kimikariek eta biologoek gehien aztertzen duten arloetako bat da. Ikerketa nabarigarrienak molekulek duten eragin zitotoxikoan oinarritzen dira ADN nuklearrarekin elkarrekintza dutenean, zelulen erreplikazio eta transkripzio makineria oztopatuz [16, 28]. ADNarekin elkarrekintza duen molekularik ezagunena, minbiziaren aurkako kimioterapia tratamenduetan erabiltzen hasi zenetik, cisplatinoa da [96]. ADNaren aurkako eraso hauek ez dituzte tumore zelulak eta zelula osasuntsuak bereizten, ondorioz, albo-ondorioak larriak aurkezten dira ehun osasuntsuetan sortzen diren kalteengatik [103, 137]. Cisplatinoaren toxikotasunarengatik, gaur egun, bere erabilera minbizi terapietan erabiltzen den azken baliabidea da, toxikotasun baxuagoa duten konposatuei ateak irekiz. Minbiziari aurre egiteko modu alternatiboetako bat molekulak ADN ezohiko egituretara bideratzea izan daiteke, hala nola G-quadruplexu ADN egitureta. Helburu berri honek cisplatinoa ordezkatzeko ikerketa berrien etorkizuna bideratzen du.

7.1 ADN G-QUADRUPLEXUA

G-quadruplexuak (GQ) guaninatan aberatsak diren azido nukleiko sekuentzietan eratzen diren ADN egitura sekundarioak dira. Hauen oinarrizko egiturari G-laukote edo G-tetrad gisa ezaguzten da eta Hoogsteen hidrogeno lotura bitartez elkarreragiten duten lau guanina basek osatutako egitura lau karratu bat da [219] (24. irudia) G-laukote egiturak pilatzean GQ osatzen da (25. irudia). Guanina baseen O6 atomoak egituraren erdigunera orientatzen dira, egituraren erdian espazio

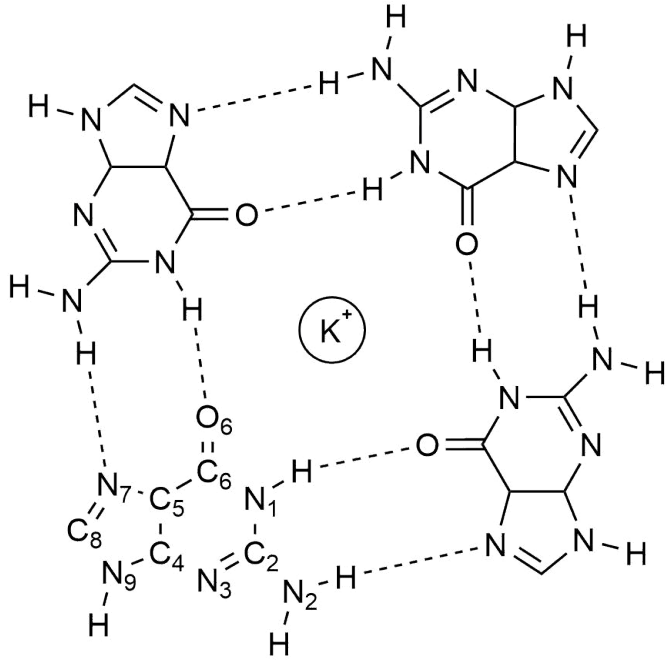

Figure 24: G-laukote egitura potasio katioi zentralarekin.

tubular huts bat sortuz, ioi-kanala deritzona. Metal alkalinoetako katioi bat ioi-kanal honetan ezartzen bada zortzi O6 guanina atomorekin elkarrekintza osatzen du, GQ egituraren egonkortzea dakarrena [109]. Gainera, plano ezberdinetako G-laukoteen artean hainbat nukleobase tartekatzen dira plano ezberdinak konektatuz. GQak hainbat modutan aztertu daitezke: monomero gisa G-laukote kopuru ezberdinekin, kate polaritatea, albo katearen topologia, edo multimero gisa azter daitezke, ehunka GQ monomerok osatutako egitura konplexu bat balitz bezala [31, 132]. GQ osatu ahal izateko, lehenik ADN helize bikoitz egitura bitan banandu behar da eta gero guaninaz ugaria den sekuentzia katea G-quadruplexua eraikitzen du. Nahiz eta GQ egitura ohikoenak kate bakarrekoak diren, lau kate ezberdinetatik eratorritako GQak ere aurkitu dira [180, 184, 204, 270].

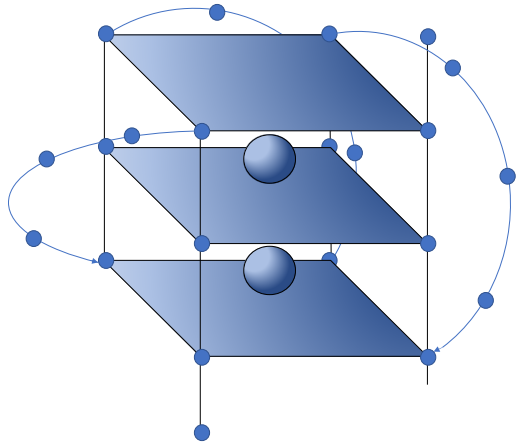

Figure 25: Hiru G-laukotez eta katioi metalikoaz osaturiko G-quadruplexua.

7.2 MINBIZIA ETA G-QUADRUPLEXUEN ARTEKO ERLAZIOA

GQak funtzio genomiko ugarirekin erlazionatuak daude, hala nola ADNaren transkripzio eta erreplikazioa prozesuekin [111], eta ikerketa ugarik agerian utzi dute GQen presentzia giza genomaren funtsezko sekuentzietan, besteak beste gene promotoreetan [228], itzuli gabeko sekuentzietan (UTR) [29], eta telomeroetan [34]. Telomeroen kasuan, GQen presentziak telomerasaren funtzioa inhibitzen duela frogatu da [71], telomeroen luzera mantentzeaz arduratzen dena eta minbizien % 85ean aktiboki parte hartzen duena [161]. Telomerasaren inhibizioak minbizi-zelulen apoptosia eragiten du eta, telomerasa minbizi-zelula gehienetan parte hartzen duenez baina zelula somatiko gutxitan adierazten denez, tumoreen kimioterapia-tratamenduetan zelula somatikoen heriotza saihestuko lukeen itu espezifiko potentzial gisa aitortzen da. Hori de eta, estrategia hau kontuan hartu da minbiziaren kontrako tratamenduetarako droga berrien ikerkuntzan, GQak egonko-

rtzen dituzten molekula txikiak aztertzen dituzten ikerketak gaur egungo aktualitate bihurtu dira [70].

7.3 ADN G-QUADRUPLEXU EGITURAREN EGONKORTZEA

GQ egitura egonkortu egin behar da telomerasaren inhibizioa eta onkogeneen adierazpena murrizten direla ziurtatzeko. Molekula organiko txiki eta lau askok [6], baita metal-konplexu batzuek ere [48], GQekin afinitate handia ez ezik, GQekiko selektibitate handiagoa duetela erakutsi dute kate bikoitzeko ADNarekin baino. Molekula hauek GQekin muturreko pilaketa (end-stacking), molekula albo kanalean tartekatzea, edo albo katearen bitartez elkarrekintzak sortu ditzakete (26. irudia). Molekula organiko lauak sakonki aztertuak izan dira, eta molekulek GQekin elkarrekintza areagotzeko beharrezkoak diren ezaugarriak zeintzuk izan behar diren frogatu da: (*a*) π-deslokalizatutako sistema bat, G-laukoteekin elkarrekintza laguntzen duena; (*b*) GQaren ioi-kanalaren O6 atomoekin elkarrekintza sortu dezakeen karga positibo bat; (*c*) positiboki kargatutako ordezkatzaileak gehituz, elektronegatiboak diren ADN katearen fosfato taldeekin elkarrekintza sortzeko gai direnak; eta (*d*) G-laukoteen antzeko azaleraduen egitura laua. Ezaugarri hauek kontuan hartuta, aztartu diren molekula gehienak eraztun aromatiko ugari dituzten molekulak dira, hala nola, porfirinak, akridinak, akridonak, kinakridinak, antrakinoak, porfirazinak, perilenoak, kinoantroxazinak, barberinak, bistriazolak edo koroneak, besteak beste. Bestalde, konplexu metalikoek propietate estruktural eta elektroniko ugari eskaintzen dituzte, molekula berriak diseinatzerakoan ustiatu daitezkeenak. Atomo metalikoari lotzen zaizkion ligandoak geometria eta orientazio zehatzetan antolatu daitezke GQarekin interakzio aproposena lortzeko. Egiturazko ezaugarriez gain, atomo metalikoen propietate elektronikoek koordinatutako ligando aromatikoetan elektroi-dentsitatea murrizten dute, elektroi gutxiko sistemen π interakzioak indartuz. Hau dela eta, ikertzen diren molekula

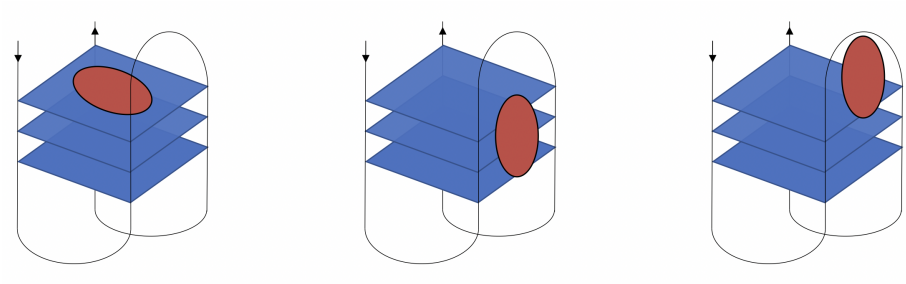

Figure 26: G-quadruplexu eta molekula txikien arteko elkarrekintza ezberdinen irudikapena. (a) Muturreko pilaketa. (b) Albo kanal lotura. (c) Albo kate lotura.

gehienak egitura organikoa duten arren, konplexu metalikoekiko interesa areagotu egin da azken urteotan.

7.4 METODO KONPUTAZIONALAK ETA EGUNGO EGOERA

Azken hamarkadetan GQekin elkarrekintza duten molekulak, hauen lotze-modua eta egonkortze mekanismoak aztertu dira, GQen eraketa eta egonkortzea arrazionalizatzeko helburuarekin. Ikerketa askotan, ikerketa esperimentalak metodo konputazionalen bidez lagundu dira, GQen eraketa eta hauek ligandoekin duten elkarrekintza hobeto ulertzeko. GQak metodo konputazionalen bitartez ikertzeko erabiltzen diren metodoak bi multzotan bana daitezke: (*a*) metodo ez kuantikoak eta (*b*) metodo kuantikoak. Lehenengoak dinamika molekularrak dira (MD), denbora luzeko simulazioak egitea ahalbidetzen dutenak, hortaz aparte metadinamikak eta ale-lodiko indar-eremu metodoak daude, oraindik simulazio denbora luzeagoak eskaintzen dituztenak. Bigarrenik metodo kuantikoak, sistema txikiagoak aztertu ditzaketen metodoak dira zehaztasun eta eraginkortasun ezberdinak eskainiz, besteak beste metodo erdi-enpirikoak (SE), mekanika kuantikoa / mekanika molekularra (QM/MM) teknikak eta dentsitate funtzionalaren teoria (DFT). Metodo hauek sistemaren egitura elektron-

ikoa esplizituki tratatu ditzazkete. Hori dela eta, sistemaren ikuspuntu estatikoagoa lortzeko erabiltzen dira, sisteman parte artzen duten elkarrekintzen egitura elektroniko zehatza deskribatzeko gai direlako. Dena den, metodo hauek mugatuak daude eskaintzen duten zehaztasuna, orokorrean, handia dela eta, konputazionalki garestiak dira eta denbora laburreko simulazioak baino ez dituzte onartzen.

GQ egiturak eta hauen prozesu biologikoak konputazionalki ikertzeko garaian metodo erabiliena dinamika molekularra da, bi arrazoi nagusirengatik. Lehenik, ADN-molekula elkarrekintza gehienak prozesu fisikoak dira, non ez diren erreakzio kimikoak gertatzen, eta jakina da erreaktibotasunaren tratamendua dinamika molekularren alderdi ahuletako bat dela. Beraz, erreakzio kimikorik ez badira gertatzen GQen egonkortze prozesuan, dinamika molekularrak erabiltzea, printzipioz, egokia da. Bigarrenik, GQ sistemak, ikerketa konputazional arloan, nahiko handiak dira, sistema biologiko gehienak bezala. Sistema mota honetarako dinamika molekular simulazioak erabiltzea ohikoa da komunitate zientifikoan, zehaztasuna galeraren truke sistema osoa aztertzeko aukera ematen baitu, sistemak denboran zehar izandako bilakaera kontuan hartuta. Dinamika molekularreko simulazioak aurrera eramateko parm99 eta parmbsco indar-eremuak, AMBER eta GROMACS softwareekin erabiltzen dira orokorrean. Edonola ere, kontu haundiz erabaki behar da erabiliko den indar-eremua, denak ez baitute errealitatea modu berdinean deskribatzen. Orokorrean, parmbsco indar-eremuak GQ sistemak hobekien deskribatzen dituen indar eremua da. Esan beharra dago dinamika molekularrak QM/MM kalkuluekin osatu daitezkeela, sistemaren zati garrantzitsuen egitura elektronikoa eta G-laukotea egonkortze prozesuan gertatzen diren interakzio eta prozesuak hobeto deskribatu eta ulertzeko. QM/MM maila goreneko metodo konputazionaltzat hartzen dira molekulek ADNarekin elkarrekintza duten sistemak aztertu nahi direnean. Zoritxarrez, aurrerapen gehiago behar dira indar-eremuak osatzen duten parametroek ez baitute GQ sistemak nahi bezain ze-

hatz deskribatzen. Šponer ikerlaria, bere taldearekin bat, aurrerapen handiak egin ditu azken hamarkadan GQ sistemak deskribatzeko erabiltzen diren indar-eremuen parametroak garatzen, batez ere tortsioekin erlazionatuak dauden parametroak [80, 98, 112, 114, 195, 230, 272]. GQen tolesdura eta osaketa prozesuak ikertzeko dinamika molekularrak dira, berriz ere, metodo erabilienak, non interakzio ahulen deskribapen egokiak zeregin erabakigarria betetzen duen. Albo-kateekin gertatzen den bezala, indar-eremu egokia aukeratzea funtsezkoa da prozesu hauen deskribapen aproposa egiteko, eta berriz ere, denbora luzeko simulazioak beharrezkoak dira, GQen tolestatze eta zabaltze prozesuak biologikoki motelak baitira mundu kuantikoaren ikuspuntutik. Hau dela eta, GPUak ez ezik, metadinamika eta ale lodiko indar-eremu metodoak askotan erabiltzen dira simulazio luzeagoak egin ahal izateko. Beste aldetik, dinamika molekularrak GQek molekula organiko zein konplexu metalikorekin duten elkarrekintza bitartez eratorritako egonkortasuna aztertzeko erabiltzen dira. Azken urteotan, ioi eta molekulek GQekin duten elkarrekintzatik eratorritako egonkortasuna ikertzen duten lanak ohikoak dira literaturan. Medikuntzaren ikuspuntutik arlo hau da interesgarriena, izan ere, lan horien emaitzak eta ondorioak erabiliz, molekula berriak sintetizatzen dira, hauek GQekin duten interakzioa eta lortzen duten egonkortze maila areagotzeko. Hau dela eta, etorkizunerako molekula organiko zein konplexu metaliko horien ikerketak oso erabilgarriak dira. Hala eta guztiz ere, kontuan hartu behar da molekula hauek GQekin ez ezik, helize bikoitzeko ADNarekin ere badutela afinitatea, hortaz, selektiboak izan behar dute baita ere, ahalik eta gehien GQen alde eginez. Helize bikoitzeko ADN eta GQek molekula ezberdinekin duten afinitatea eta selektibitatea aipagarriak dira arlo honetan Barone eta bere taldearen lan konputazionalak [36, 37, 85, 240, 241] eta Neidle-en taldearen lan esperimentalak [6, 56, 90–92, 113, 180].

GQekin punta-puntako ikerketa konputazionalak ez dira dinamika molekularretan eta QM/MM simulazioetara mugatzen, eta dispert-

sioaren bidez zuzendutako dentsitatea funtzio teoria (DFT-D) kalkuluak ere aurkitu daitezke bibliografian atomo kopuru murriztuagoko sistemetan. DFT-D metodoak egitura elektronikoa esplizituki deskribatu dezake, molekula eta GQek arteko elkarrekintza modu zehatzenean deskribatuz. Zoritxarrez, metodo hauek ezin dute atomoek denboran zehar duten bilakaera ikertu. DFT-D metodoak Fonseca Guerra eta bere lankideek GQekin zerikusia duten ikerketa askotan erabili dute [88, 89, 260, 266, 267], batez ere, G-laukote kopuru ezberdineko sistematan, ioi-kanalean kokatzen diren atomo metalikoek duten garrantzia, G-laukotea osatzen duten nukleobaseen aldaketak (adenina bat guanina batengatik aldatzea adibidez) sortu dezakeen egonkortasun edo desegonkortasuna , eta elkarrekintza ahulak (hidrogeno loturak eta $\pi - \pi$ elkarrekintzak) deskribatzeko, besteak beste. DFT-D metodoen bitartez eratorritako ondorioetako bat katioi alkalinoak ez direla derrigorrezkoak GQak egonkortzeko da baina, G-laukoteetako O6 atomoekin duten interakzioa ioi-kanalean ezartzean GQen egonkortasuna areagotu dezakete. DFT-D metodoen bitartez lortutako beste ondorio interesgarri bat RNAn osatzen diren GQak ADNan osatutakoak baino egonkorragoak direla datza, RNAren erribosaren 2-OH fosfato taldeen oxigeno atomoekin erlazionatuta dagoen hidrogeno-lotura gehigarri batekin erlazionatuta dagoela dirudi, egiturari konformazio-egonkortasun handiagoa ematen diona. Azkenik, jakina da, dispertsio zuzenketak metodo erdi-enpirikoetan gehitu aurretik zaila zela metodo hauek sistema biologikoetan handietan aplikatzea. Zoritxarrez, ikertzaile gutxi saiatu dira metodo erdi-enpirikoen bitartez GQak aztertzen, eta orokorrean metodo erdi-enpirikoak GQen kargak kalkulatzeko erabiltzen dira gero beste metodo zehatzago batean erabiltzeko. Hala ere, badaude lan interesgarriak, adibidez, Dinçalp et al. egindakoa [60], non propietate espektroskopikoak aztertzeko metodo erdi-enpirikoak erabili ziren.

7.5 ESKALATZE-LINEALEKO METODO KONPUTAZIONALAK

Azken urtetan eman diren ordenagailuen gaitasun hazkuntza eta konputazio tekniken garapenak sistema molekular handiagoen ikerkuntza ahalbidetu du. Eskalatze-linealeko metodoak gaur egungo egoerarekin bat dator ordenagailuek paralelizazio masiboaren bitartez eskaintzen duten gaitasuna maximizatuz, sistema molekular handiko kalkuluak denbora tarte laburrean burutzen ahalbidetuz. Metodo berri hauek, eskaintzen duten abantailengatik, kimika konputazionalaren etorkizunaren nondik norakoa zehazten dute. Eskalatze-linealeko metodo gehienak gaur egun dauden DFT metodoen algoritmoak berritzean oinarritzen dira. Metodo hauek, sistema molekular handiak ez ezik, sistema osoaren egitura elektronikoa deskribatu dezakete. Nahiz eta ikerketak badauden, eskalatze-linealeko DFT metodoak tamaina handiko sistema biologikoak ikertzeko gutxitan erabiliak izan dira [178, 179, 227]. Beste aldetik, eskalatze-linealeko kluster akoplatu (CC) metodo berri bat argitaratu da, DLPNO-CCSD(T) deiturikoa. Metodo berri hau nabarmenki eraginkorra da eta CC metodoen aplikazioa ahalbidetzen du sistema biologiko handietan [199, 200]. Kontuan hartu behar da CC metodoak sistema molekular oso txikietara zeudela bideratuak metodo berri hau argitaratu aurretik.

7.6 TESI HONEN HELBURUA

GQak azken hamarkadan interes handiko gai bihurtu dira. ADN egitura sekundario hauekin zerikusia duten mekanismo biologikoak eta hauen interakzioa molekulekin ikertu eta ulertzea berebiziko garrantzia hartu du, egungo minbiziaren kontrako tratamendu kaltegarrien hautabidea eskaintzen baitute. Honako tesi honetan GQak eta hauen ikerkuntza metodo konputazionalen bitartez lantzen da. Tesi honek hurrengo gaiak jorratzen ditu:

- Metodo erdi-enpiriko, eskalatze-linealeko DFT, eta hurbileko eskalatze-linealeko DLPNO-CCSD(T) metodoen erabilera G-quadruplexu egiturak aztertzeko, eta hauek molekulekin duten interakzioa deskribatzeko.

- G-quadruplexu eta [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] konplexu metalikoaren arteko interakzioa aztertu eta deskribatu. Ikertuko den konplexu metalikoa dagoeneko sintetizatua izan da eta tumore mota ezberdinetan eragina duela frogatu da. Konplexu metaliko honen eraginkortasuna nondik datorren landuko da.

- [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] konplexu metalikoak helize bikoitzeko ADN eta G-quadruplexu ADN egiturekin duen afinitatea eta selektibitatea eztabaidatuko da.

8

METODOAK

8.1 TEORIA KUANTIKOA

Kimika konputazionala eta modelatze molekularra molekulen portaera ezaugarritzeko eta iragartzeko erabiltzen dira. Sistema molekularren energia kalkulatzeko metodorik zehatzena Schrödingerren ekuazioa ebaztea da, hau da, hain zuzen ere, mekanika kuantikoaren oinarria. Fisikaren teoriak, joan den mendeko hogeiko hamarkadan garatuak [61, 101, 215, 216], fisikarien mundu mikroskopikoaren ikuspuntua aldatu zuen. Bohrrek atomo sinpleenaren egonkortasuna azaldu zuen, hidrogenoa, elektroiak nukleoaren inguruan orbita egonkorretan zeudela finkatuz. Nahiz eta hasieran teoriak besterik ez ziren laster aurkitu ziren aplikazioak kimikan, kimika kuantikoa deritzona sortuz. Kimika kuantikoan molekula baten elektroiak esplizituki tratatzen dira, eta hori informatikoki tratatzea garestia da; azken horrek esan nahi du metodoa zoritxarrez atomo kopuru txikiko sistemetara mugatu izan dela. Urteetan zehar, kimika kuantikoak tresna eta metodoak ezberdinak sortu eta garatu ditu propietate molekularrak kalkulatu, ulertu eta aurreikusteko.

Molekula handientzat, DNA adibidez, sistemaren deskribapenak burutzeko metodo "sinpleagoak" erabili behar ditu, hala nola MM. Metodo hauek mugimenduaren printzipio klasikoetan oinarritzen dira, atomoen energia posizioarekiko funtzioa bezala ebaluatuz. Elektroiak ez dira esplizituki tratatzen metodo hauetan.

Mekanika kuantikoan, printzipioz, sistema edo molekula baten informazio guztia, *uhin-funtzio* bidez lor daiteke, Ψ. Hori zehazteko,

Schrödingerren ekuazioa baino ez dugu ebatzi behar (hemen bere denborarereiko independente egituran):

$$\hat{H}\Psi = E\Psi \tag{8.1}$$

Hamiltondarrak, $\hat{H}$, elektroien eta nukleoen energia zinetikoa, nukleo-elektroi arteko elkarreraginak, eta nukleo-nukleo eta elektroi-elektroi arteko errepultsioa deskribatzen dituen. $\hat{H}$ uhin-funtzioari aplikatzean, sistemaren energia, E, bider uhin-funtzioa lortzen da. N_e elektroiak eta N_N nukleoak dituen sistema batentzat, unitate atomikoetan Hamiltondarra 8.2 ekuazioan bezala adieraziko litzateke, efektu erlatibistak eta espin-orbita akoplamendua alde batera utzita.

$$\begin{aligned} \hat{H} &= \hat{T}_e + \hat{T}_N + \hat{V}_{e-e} + \hat{V}_{e-N} + \hat{V}_{N-N} \\ &= -\sum_i^{N_e} \frac{\nabla_i^2}{2} - \sum_A^{N_N} \frac{\nabla_i^2}{2M_A} \\ &\quad -\sum_A^{N_N}\sum_i^{N_e} \frac{Z_A}{r_{iA}} + \sum_{i>j}^{N_e} \frac{1}{r_{ij}} + \sum_{A>B}^{N_N} \frac{Z_A Z_B}{R_{AB}} \end{aligned} \tag{8.2}$$

8.2 ekuazioan $\hat{T}_e$ eta $\hat{T}_N$ elektroien eta nukleoen energia zinetikoari dagokio, hurrenez hurren; $\hat{V}_{e-e}$, $\hat{V}_{e-N}$, eta $\hat{V}_{N-N}$ terminoek, hurrenez hurren, elektroi-elektroi, elektroi-nukleo eta nukleo-nukleo elkarrekintzak adierazten dituzte. M_A, A nukleoaren masa da, Z_A eta Z_B, A eta B nukleoen zenbaki atomikoak dira, hurrenez hurren; eta $R_{AB} = \left|\vec{R}_B - \vec{R}_A\right|$, $r_{ij} = \left|\vec{r}_j - \vec{r}_i\right|$, $r_{iA} = \left|\vec{R}_A - \vec{r}_i\right|$ A eta B nukleoen, i elektroi eta j elektroien, eta i elektroien eta A nukleoen arteko distantziak dira, hurrenez hurren.

8.2 ekuazioa Born-Oppenheimerren hurbilketa ezagunarekin sinplifikatu daiteke [38]. Nukleoak elektroiak baino astunagoak direnez, euren mugimendua motelagoa izango da, horregatik nukleoaren termino zinetikoa, $\hat{T}_N$, alde batera utzi daiteke. Ondorioz, nukleo-nukleo potentzial energia terminoa konstante gisa erabil daiteke, $\hat{V}_{N-N}$ nuk-

leoen arteko distantzien araberakoa baino ez delako, horiek finkoak baitira. Born-Oppenheimer hurbilketa 8.3 ekuazioan agertzen den hurrengo Hamiltondarra adierazten du.

$$\begin{aligned} \hat{H}_e &= \hat{T}_e + \hat{V}_{e-e} + \hat{V}_{e-N} + \hat{V}_{N-N} \\ &= -\sum_i^{N_e} \frac{\nabla_i^2}{2} - \sum_A^{N_N} \sum_i^{N_e} \frac{Z_A}{r_{iA}} + \sum_{i>j}^{N_e} \frac{1}{r_{ij}} + \sum_{A>B}^{N_N} \frac{Z_A Z_B}{R_{AB}} \end{aligned} \tag{8.3}$$

$\hat{H}_e$ Hamiltondar operadore elektronikoa da. Uhin-funtzio elektronikoa, Ψ_e, esplizituki elektroien koordenatuen araberakoa da, eta parametrikoki nukleoen posizioaren araberakoa. Aipatzekoa da edozein nukleoen kokapenerako energia elektronikoa lortzen dela, baita nukleo-nukleo elkarrekintzarena ere. Honek $3N_N$-dimentsioko energia potentzial eraginkor bat eraikitzen du (N_N nukleoentzat), honen topologiak konformazio ezberdinen energiak eta energetikoki lokalak diren bi egitura minimo lotzeko bideak emango ditu, askotan Energia Potentzial Gainazala (PES) deitzen dena. PESak atomo multzo baten beharrezko informazioa ematen duen ezaugarria da. Transformazio kimikoen mekanismoak eta horien energiak bereizteko behar den informazioa ematen du.

Metodo matematiko asko garatu dira edozein antolamendu nuklearri uhin-funtzio elektronikoa aurkitzeko modua emateko. Garatu zen lehen metodoetako bat Hartree-Fock metodoa (HF) izan zen [72, 97], non uhin-funtzioa elektroi bakarreko orbital antisimetrikoen produktu bat den. Elektroiak potentzial eremu batean mugitzen dira inguruko nukleoen eta beste elektroien ondorioz. Elektroien arteko korrelazioa kontuan hartzeko modu ezberdinak daude. Horietako bat n^{th} ordenako Møller-Plesset teoria (MPn) bezalako perturbazio-metodoei dagokie [164]. Metodo honetan elektroi-korrelazioa HF arazoaren perturbazio gisa tratatzen da. Ezarpen Interakzioa metodoan (CI) uhin-funtzioa Slater determinatzaileen konbinazio lineal gisa adierazten da [221, 222], elektroi ugariko uhin-funtzio zehatzei bariazio-soluzio

hobea emateko. Metodo sofistikatuagoak daude, hala nola Kluster Akoplatua (CC) [24–26, 186], Espazio-Aktibo-Osoko Eremu Autokerentea (CASSCF) [3, 4] edo Elkarrekintza Anitzeko Konfigurazio Erreferentzia (MRCI) [130, 258]. Metodo hauek tresna baliagarriak eta zehatzak dira oinarrizko zein kitzikatutako egoeren propietate elektronikoak aztertzeko.

Badira beste metodo batzuk, hala nola metodo erdi-enpirikoak (SE), askoz azkarragoak direnak, baina ez hain zehatzak, sakoneko fisikaren tratamenduan hurbilketa asko eta datu enpirikoen parametroak dakartzatelako. Hurbilketa zehatzagoa DFT metodoa da [105, 131, 181, 182], HF metodoarekin (*ab initio* metodo sinpleena) oso erlazionatua dagoena, baina korrelazio elektronikoa kontuan hartzeko modua ezberdina duena, bere zehaztasuna asko hobetzen duena.

8.2 METODO ERDI-ENPIRIKOAK

Metodoa erdi-enpirikoa (SE) HF metodoaren sinplifikazioa da, errendimendua handitzeko helburuarekin. SE metodoetan balentzia-elektroiak bakarrik hartzen dira kontuan esplizituki, nukleo-elektroiak kontuan hartzen dira nukleoek eta elektroiek eragiten duten errepultsio konbinatua modelatzeko funtzioak sartuz. Gainera, *Slater motako orbital* (STO) esponentzialen gutxieneko funtzio kopurua erabiltzen da oinarri gisa balentzia-elektroiak egokitzeko.

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_l^m(\theta,\varphi)r^{n-l}e^{-\zeta r} \tag{8.4}$$

non N normalizazio konstante bat den, Y_l^m funtzio harmoniko esferikoak diren, *n*, *m* eta *l* zenbaki kuantikoak diren, θ angelu polarra den eta φ angelu azimutala den, r nukleotik elektroiak duen distantzia den, eta ζ nukleoaren karga eraginkorrarekin erlazionaturik dagoen konstantea den.

SE metodoetako askok Zero-Diferentzial Gainjartze Hurbilketa (ZDO) erabiltzen dute. Hurbilketa honetan elektroien arteko errepultsio integralak baztertu egiten dira. Hiru- edo lau-zentrotako kargak banatzen dituzten bi elektroi integral guztiak alde batera uzten direla, eta gainerako integralak parametroen bitartez kontuan artzen direla, datu esperimentalak edo maila altuko kalkuluak erreproduzitzeko helburuarekin. Horregatik deitzen dira metodo hauek erdi-enpirikoak. SEetan metodo desberdinak bereizten dira zenbat integral baztertzen diren eta parametrizazioa nola egiten den kontuan hartuta, hala nola Diatomoen Gainjartze Diferentzialaren Baztertzea (NDDO), Diatomoen Gainjartze Diferentzialaren Bitarteko Baztertzea (INDO) edo Gainjartze Diferentzialaren Erabateko Baztertzea (CNDO). Lan honetan, NDDO modeloa eta bere bertsio hobetuak erabili dira.

8.2.1 *Diatomoen Gainjartze Diferentzialaren Baztertzea (NDDO)*

Hurbilketa honetan, nukleo ezberdinetan dauden bi orbital atomikoetatik abiatuz, termino guztiak zero dira, eta, beraz, honako gainjartze integrala (8.5 ekuazioa) eta elektroi bakarreko operadorea (8.6 ekuazioa) sortzen dira:

$$S_{\mu\upsilon} = \langle \mu_A | \upsilon_B \rangle = \delta_{\mu\upsilon}\delta_{AB} \tag{8.5}$$

$$\hat{h} = -\frac{1}{2}\nabla^2 - \sum_{A}^{N_N} \frac{Z'_A}{\left|\vec{R_A} - \vec{r}\right|} = \frac{1}{2}\nabla^2 - \sum_{A}^{N_N} V_A \tag{8.6}$$

Eraldatutako metodoek datu esperimental molekularrak erabili zituzten parametroak egokitzeko. MNDO, AM1 eta PM metodoak NDDOren hurbilketatik eratorriak dira, eta nukleo-nukleo errepult-

sioa tratatzeko moduan eta parametroak esleitzeko moduan bereizten dira.

8.2.2 *Metodo Parametrikoen Ereduak (PM)*

AM1 garatu ondoren, parametroak optimizatzeko metodoa hobetu zen. Horren emaitza Metodo Parametrikoa 3 (PM3) izan zen, Stewartek garatua [231–233]. PM3 metodoan, nukleo-nukleo errepultsioa AM1-ek duen adierazpena mantendu zen baina, kasu honetan, bi nukleo-nukleo gausstar funtzio esleitu zitzaizkion atomo bakoitzari (8.7 ekuazioa).

$$E_N^{MNDO}(A,B) = E_N^{MNDO}(A,B) + \frac{Z'_A Z'_B}{R_{AB}} \times \left[\sum_k \alpha_{kA} e^{-b_{kA}(R_{AB}-c_{kA})^2} + \sum_k \alpha_{kB} e^{-b_{kB}(R_{AB}-c_{kB})^2}\right] \tag{8.7}$$

Gainera, parametrizazio estrategia desberdina zen. AM1 batez ere, datu atomiko kopuru txikia erabiliz parametrizatu zen, baina PM3 propietate molekular ugari erreproduzitzeko parametrizatu zen. PM3 ondo deskribatzen ditu hidrogeno loturak, baina hidrogeno-hidrogeno lotura erakargarri ez-fisikoak ganbaloratzen ditu kasu askotan, arazo larriak sortzen duena molekulen arteko elkarreraginak edo molekula malguen konformazioak aztertzean. Gainera, *s*- eta *p*-funtzioak bakarrik kontuan hartuta, metodoak ezin du taula periodikoaren zati handi bat deskribatu. Jakina da *d*-orbitalek nabarmen hobetzen dituztela taula periodikoaren bigarren lerroko elementuentzako emaitzak, bereziki hiperbalenteak izateko potentziala dutenenak. Orbitalak kontuan hartzearen arazo nagusia bi-elektroi integralen gehikuntzan sortzen da. Hala ere, Thielek eta Voityukek frogatu zuten *d*-orbitalak gehitzean lortutako zehaztasun handiak [242–244] kostu konputazionala gainditzen duela.

NDDO nukleo-nukleo interakzio terminoaren zenbait aldaketaren ondoren, parametroen optimizazioaren metodoak eta *d*-orbitalen gehikuntzak, PM6 metodoaren sorkuntza eragin zuten. PM6 metodoan erabiltzen den nukleo-nukleo elkarrekintzaren forma orokorra 8.8 ekuazioak ematen du.

$$E_N^{PM6}(A,B) = Z'_A Z'_B \langle s_A s_A | s_B s_B \rangle \left[1 + \chi_{AB} e^{-\alpha_A (R_{AB} + 0.0003 R_{AB}^6)}\right] \quad (8.8)$$

SE metodoeak hauek zuten ahulezietan egin ziren hurrengo hobekuntzak, elkarreragin ez-kobalenteetan eta bereziki hidrogeno loturetan. Hainbat zuzenketa egin zitzaizkion PM6 metodoari, eta bertsio desberdinak sortu ziren, hala nola PM6-D3 [84], PM6-DH [196], PM6-DH+ [133], PM6-DH2 [135], PM6-D3H4 [198], PM6-DH2X [197], eta PM6-D3H4X [41]. Zuzenketak bi taldetan bana daitezke: Dispertsio-zuzenketak eta Hidrogeno-lotura zuzenketak. Elkarreragin horiek ezinbestekoak dira zehaztasun kimikoa deritzona lortzeko.

DFT-D metodoan garatutako Grimme D3 sakabanatze-zuzenketa PM6 metodoari aplikatu zitzaion, zein arrakastaz probatu zen milaka sistema ezberdinetan, inter- eta intra-elkarrekintza molekularrak kontuan artuta.

Bestalde, PM6 metodoak jasan duen beste zuzenketa garrantzitsua hidrogeno-loturan dago, hidrogenoak lotzeko hiru zuzenketa-belaunaldi ezberdinetan bereizi baitira.

Lehen belaunaldiko zuzenketan [135] kargak (q) erabili ziren atomo onartzaileetan (A) eta hidrogeno atomoetan (H), H-lotura distantzia (r) eta kosinu terminoa A-H-D angelurako 180º-ko egoera sustatzen zuena. Zuzenketa honen arazo nagusia H-lotura distantziaren mende bakarrik zegoela izan zen, A-H-D angelurako 90º inguruko potentzial etenak eragiten zituena. Akats horiek bigarren belaunaldiko zuzenketetan konpondu ziren ziren.

Bigarren belaunaldiko H-lotura zuzenketak forma hau zuen [135]:

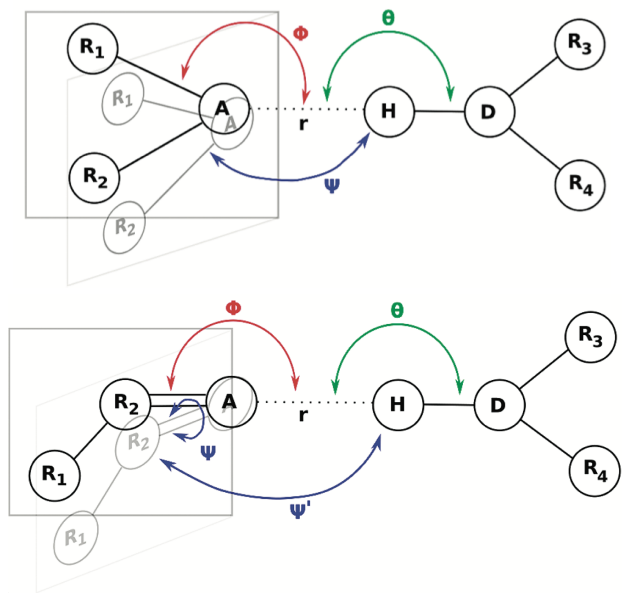

Figure 27: Hidrogeno loturaren ezaugarri geometrikoen ilustrazioa bigarren belaunaldiko zuzenketetan.

$$E_{H-bond} = \left[a\frac{q_A q_H}{r^b} + cd^r\right]\cos(\theta)\cos(\phi)\cos(\psi) \tag{8.9}$$

Zuzenketa hau PM6-DH2 metodoan gauzatu zen, doktoretza tesi honetan erabilia izan dena. Sei barne-koordenadarekin definitu daiteke (ikusi 27 irudia). Horrekin, bigarren belaunaldiko bertsioak fisikoki desegokiak diren ekuazio terminoak eta parametroak saihesten ditu, metodo sendoago bat sortuz.

Hirugarren belaunaldiko zuzenketek [133] H-lotura bi atomoren arteko atomo-atomo terminoa karga hartzaile edo emaileari independentetzat hartzen dute, f_{geom} (8.10 ekuazioa) funtzio batez haztatua. Honek bi fragmentuen arteko orientazio erlatiboaren araberakoa da, gainera, funtzio izurtzaile batez biderkatua, distantzia laburrak eta luzeak zuzentzeko.

$$f_{geom} = \cos(\theta_A)^2\cos(\phi_A)^2\cos(\psi_A)^2\cos(\phi_B)^2\cos(\psi_B)^2 f_{bond} \quad (8.10)$$

PM metodoen azken hobekuntzak PM7 metodoan bildu ziren [235]. Hamiltondarra birparemetrizatu zen zehaztasun altuko *ab initio* metodoak eta datu esperimentalak erabiliz eta, gainera, bi aldaketa egin ziren metodoan: elkarrekintza ez-kobalenteen deskribapena hobetzeko aldaketa bat, eta NDDO formalismoaren bi akats txiki zuzendu ziren. Bere aurrekoaren akats batzuk finkoak izan ziren, hala nola atomo bikote batzuen arteko gaitzespen murriztua edo desagertua: Na-Na, Br-N, Br-O, Br-Br, S-N, S-S, S-O, S-Cl, I-N, I-O eta I-I. Horrez gain, aldaketa batzuk kontuan hartu ziren [234], nukleo-nukleo elkarrekintzaren eta elektroi-elektroi errepultsio integralaren balioari muga bat gehituz, termino batzuek balio zehatzetan bat egiteko aukera izan dezaten 7 Å-tik gorako banaketetan, egoera solidoaren tratamendu hobea ahalbidetzeko.

8.3 DENTSITATE FUNTZIONALAREN TEORIA METODOAK

Dentsitatearen Teoria Funtzionalaren formalismoak N-elektroi uhin-funtzioa eta Schrödingerren ekuazioa hiru aldagai espazialeko funtzioa den elektroi dentsitatearengatik ordezkatuz, $\rho(\overrightarrow{r})$, askoz sinpleagoa ebazten dena. Ondoren, edozein sistemaren egoera elektronikoa, energia eta propietate elektroniko guztiak deskribatu daitezke $\rho(\overrightarrow{r})$ ebaluatuz [66, 182].

Hohenbergek eta Kohnek frogatu zuten oinarrizko egoera ez endekatua duen sistema baten propietate elektronikoak, elektroi dentsitateak determinatzen dituela. Horregatik, oinarrizko energia, E_o, $\rho(\overrightarrow{r})$-ren determinatzen du. Energia-aldakuntzako printzipio bat ere ezarri zuten energia-funtzionalerako, uhin-funtzioetarako aldakuntza-printzipioaren antzekoa. Horrela, $E[\rho]$ funtzioaren forma zehatza

ezagututa, oinarrizko egoeraren dentsitatea bila dezakegu (uhinaren funtzioaren kasuan bezala). Kohn-Sham DFT bariazioan [131] N-elektroi molekula baten E_o energia elektronikoa, ρ probabilitate dentsitatea duena, hurrengo ekuazioaren bidez ematen da:

$$\begin{aligned} E_0 &= -\tfrac{1}{2}\sum_{i=1}^{N}\left\langle \psi_i(1) \mid \nabla_i^2 \mid \psi_i(1)\right\rangle + \int v(r)\rho(1)d\vec{r_1} \\ &\quad +\tfrac{1}{2}\int\int \frac{\rho(1)\rho(2)}{r_{12}} d\vec{r_1}\, d\vec{r_2} + E_{xc}[\rho] \end{aligned} \tag{8.11}$$

non $v(r) = -\sum_\alpha \frac{Z_\alpha}{r_{1\alpha}}$ nukleoek sortutako kanpo potentziala, ψ_i Khon-Sham orbitalak, eta E_{xc} trukatze-korrelazioaren energia diren.

Dentsitate funtzionaleko metodoek emaitza bikainak eman dituzte sistema kimiko gehienetan [140]. Hala ere, sarritan loturen disoziatzeko energiak gainbaloratzen dituzte[250]. HF metodo hibridoak eta DFT teoriek energia disoziatiboen zehaztasuna areagotzen dute, Johnsonek *et al.* balioztatu zuen bezala [116].

Metodo hauen abantaila nagusietako bat HF metodoen antzeko kostu konputazionalarekin, elektroi korrelazio motaren bat duzula, eragozpena batekin: korrelazio efektuak ezin direla zehazki ordenatu izanik. Dagoeneko hasieratik nahasten dira korrelaziorik gabeko soluzioarekin. Egiaztatu da DFTk emaitza onak ematen dituela hainbat sistema kimikoren propietateei dagokienez, MP2-ren emaitzen pareko kalitatearekin [271] edo are hobeagoa kasu batzuetan. Kostu "baxua" dutenez, sistema handietarako DFT hautazko metodoa da, eta, horretarako, debekatuta dago elektroi-korrelazioa sartzea MP edo CI metodoen bidez.

8.3.1 *Eskalatze-Linealeko Dentsitatearen Teoria Funtzionala*

DFT metodoak sistema solido edo biologikoetan aplikatzean duten kostu konputazionala da oztopo nagusia, sistemaren tamainua areagotzean kostuak esponentzialki handitzen baita.

Horrelako sistemen kasuan, milaka atomo dituzten sistemen deskribapen zehatza behar da. Atomo kopuru horretara iritsitako sistemek DFT kalkulu estandarren kostua azkar handitzen da, $\mathcal{O}(N^3)$ eran, non N atomo kopurua den. Arazo hau gainditzeko, hainbat metodo garatu dira, hala nola eskalatze-lineala, $\mathcal{O}(N)$ izenez ere ezagutzen dena [81]. Gainera, ordenagailuen kalkuko gaitasun hazkundeak, batez ere, prozesadore edo nukleo kopuruaren igoerarekin dator, derrigorrez, paralelizazio eraginkor bat eskatzen duena. Oro har, eskala-handiko eta linealeko metodoek funtsezko bi puntu gainditu behar dituzte sistema handietan eta masiboki paralelizatutako ordenagailuetan DFT kalkulu eraginkorrak egiteko:

- Egitura elektronikoa kalkulatzeko metodo bat garatzea, masiboki paraleloak diren kalkuluetarako egokia dena.
- Egitura elektronikoak $\mathcal{O}(N^3)$ baino eskalatze hobearekin erresolbatzea, egokiena, eskalatze linealarekin izanik, $\mathcal{O}(N)$.

Helburu horiek lortzeko hainbat modu daude: diagonalizazio iteratiboaren teknika erabiltzea [138, 185], Fourierren transformazio azkarrak (FFT) [55, 173], *ab initio* psuedopotentzialak erabiltzea nukleoko elektroiak ordezkatzeko [10, 192, 249, 254], edo uhinen-planoen kasuan, hauen kopurua murriztea uhin-planu hazituak erabiliz nukleoko eskualdeetan, uhin-planu lineal hazituetan bezala (LAPW) [35, 156, 225, 226], edo uhin areagotuen proiektorea (PAW) [104, 139]. Eskalatze-lineako edo eskala-handiko funtzionaltasunaz jabetzen diren kode nagusiak hurrengo zerrendan azaltzen dira, orden alfabetikoan: CONQUEST [46], ErgoSCF [63], FEMTECK [251, 252], FreeON [67], ONETEP [174], OpenMX [175], PROFESS [110], eta SIESTA [207].

8.3.2 *SIESTA Metodoa*

Lan honetan SIESTA izan da gure datu-iturri nagusia. SIESTA proiektua (Spanish Inititive for Electronic Simulations with Thousands of Atoms) [7, 176, 177, 213, 229] 1995ean hasi zen, eta oso eskuragarri dagoen eskalatze-lineako koderik goiztiarrenetako bat da, DFT erabat autokoherentea, oinarri orbital atomikoaren (LCAO) konbinazio lineal batean oinarritua. Born-Oppenheimer hurbilketaz gain, hurbilketarik oinarrizkoenak honako hauei buruzkoak dira: trukatze eta korrelazio tratamenduari (XC), arau-kontserbatzaile estandarrak erabiltzea, pseudopotentzialak [11, 93] beren forma erabat lokalean mantenduz [129], eta orbital atomiko numerikoen erabilera, hertsiki lokalizatuak eta erradio (r_c) jakin batetik haratago desagertzen direnak.

Horrek, printzipioz, LS-DFT eskalatzea ahalbidetzen du, metodo orokorra eta malgua. Hala ere, esan beharra dago matrize elementuak eskalatze-linealeko algoritmo batekin konputatzen badira ere, diagonalizazioa $\mathcal{O}(N^3)$-rekiko porportazionala dela.

8.3.2.1 *Pseudopotentialak*

Edozein sistemaren egitura elektronikoa ebazterakoan, elektroi guztiak kontuan hartu behar dira, beste partikula batzuek nabartzen duten potentzialean parte hartzen baitute. Praktikan, atomo baten nukleo-elektroiak ahulki perturbatzen dira aldaketa kimikoen ondorioz, balentziazko elektroiekin alderatuz, eta, beraz, hainbat hurbilketa metodo garatu dira nukleo-egoera horiek tratatzeko, gastu konputazionala murriztuz. Nukleo-elektroiak eta protoiak, aurkako kargak dituzte, partzialki elkar ezeztatzen direnak, potentzial eraginkor konbinatu batekin ordezka daitezke, pseudopotentzial bezala ezagutzen dena. Pseudopotentzialak nukleoa eta nukleo-elektroien potentzial zehatza ordezkatzen du, atomikoaren zentroarekiko erradio jakin baten barruan, potentzial eraginkor batekin. Distantzia honen

barruan, nukleoaren erradioa bezala ezagutzen dena, potentziala leundu egiten da eta nukleoan balio mugatu batera jotzen du, mugan dagoen benetako potentziarekin bat egiten duen bitartean. Potentziala leuntzearen ondorioz, balentzia-estatuen nodo erradialak galdu egiten dira erdigunean ez baita beharrezkoa nukleo-estatueen ortogonaltasuna mantentzea. Kasu gehienetan, pseudopotentzial ez-lokal bat erabiltzen da, horrek esan nahi du potentzial ezberdina dagoela momentuko kanal angeluar bakoitzarentzat, nukleoaren erradioa bereizia duena, r^l_{core}, kanal horretarako egokia.

SIESTAk arau-kontserbatzaile printzipio estandarrak jarraitzen ditu: pseudopotentzialak kontserbatzen ditu forma erdi-lokalean [93], eta $V_l(r)$ potentzial erradial desberdina du momentu angeluar bakoitzeko, erlatibistikoki sortua [11, 128]. Oro har, Troullier-Martins parametrizazioa pseudopotentziala leuntzeko erabiltzen da, behar diren baliabide konputazionalak murriztuz [249]. Forma erdi-lokal hori Kleinmanek eta Bylanderrek (KB) proposaturiko forma ez-lokal bihurtzen da [129].

8.3.2.2 *Oinarri-Funtzioak*

Kohn-Sham ekuazioen zenbakizko soluzioa orbitalak konputazionalki funtzio matematiko egoki bati dagokionez zabalduz egiten da: oinarri funtzioak. Funtzio hauek zenbaterainoko ekarpena egiten duten zehazten duten koefizienteak printzipio aldakorra aplikatuz aurkitzen dira. Kalkulu lineal handietarako oinarri-funtzio egokia aukeratzean, espazio errealean kokatu beharra eta zenbakizko doitasun arrazoizkoa lortzeko behar diren oinarrizko funtzioen kopurua minimizatzeko beharra ditugu gidari.

SIESTA metodologian, oinarri-funtzioen aukeraketa estandarra orbital pseudoatomikoak (PAO) dira, φ^A_{nml} I atomoarentzat, momentu angeluar bakoitzeko sare erradial logaritmiko batean taularatzen direnak eta gero hurbilketa harmoniko esferikoarekin biderkatzen direnak:

$$\varphi^{I}_{lmn}(r,\theta,\varphi) = R^{I}_{nl}(r)Y_{lm}(\theta,\varphi) \tag{8.12}$$

Goiko PAOak urruntasunarekin azkar gainbeheratzen diren bitartean, atomikoki zentratutako beste oinarrizko funtzio batzuk bezala, asintotikoki zerora bakarrik infinituan jotzen dute. Eskalatze-lineala lortzeko, Hamiltondarra benetako espazioan zorrozki ezarri behar da. Sankey eta Niklewski metodoari jarraituz [214], arazo pseudoatomikoaren autofuntzioak muga esferiko baten barruan aurkitzen dira, non potentziala infinitu bihurtzen den. Horrela, PAOen "isatsak" aldatu egiten dira, eta zorrotz joaten dira zerora erradio jakin batean. Erradio hau, r_c, desberdina izan daiteke momentu angeluar bakoitzeko.

8.3.2.3 *LMKLL van der Waals dentsitate funtzionala*

Van der Waalsen erakarpena (vdW) mekanika kuantikoaren fenomeno bat da, eta indar hori ulertzea eta deskribatzea funtsezkoa da egitura biologikoak aztertzeko (adibidez, DNA eta proteinen egiturak). Vd-Wren indarrak eskualde ezberdinetan dauden kargen mende daude, hortaz korrelazio efektu ez-lokala da. Tesi honetan vdW LMKLL funtzionala erabili da, dispertsioa inplizituki tratatzen duena, Grimme sakabanaketaren ohiko zuzenketak ez bezala, geometriaren menpekoak baitira. VdW-DF metodoaren gakoa korrelazio-energiaren irismen luzeko pieza bat sartzea da, ρ dentsitatearen funtzio erabat ez-lokala (8.13 ekuazioa).

$$E_c^{nl}[\rho] = \int d^3r \int d^3r' \rho(r)\phi(r,r')\rho(r') \tag{8.13}$$

8.4 DOMAINETAN OINARRITUTAKO TOKIKO ORBITAL PAREA KLUSTER AKOPLATUA KITZIKAPEN BAKAR, BIKOITZ, ETA HIRUKOITZ PERTURBATIBOAREKIN (DLPNO-CCSD(T))

Kluster akoplatuaren metodoa (CC) hurbiltze-eskema bat da elektroientzako Schrödinger ekuazio independentea ebazteko. Čížek eta Paldusek kimika kuantikoan sartu zutenetik askok erabiltzen dute [54], zehaztasun handiko kalkuluak egiteko.

Kluster akoplatuko teoriak Schrödinger ekuazio independentearen soluzio zehatza ematen du. Kluster akoplatu teoriaren uhin-funtzioa ansatz esponentzial gisa idatzita dago:

$$|\Psi\rangle = e^{T} |\Phi_0\rangle = \Omega |\Phi_0\rangle \tag{8.14}$$

non $|\Phi_0\rangle$ erreferentziazko uhin-funtzio den, normalean Hartree-Fock orbital molekularretatik (MO) abiatuta eraikitako Slater motako determinatzaile bat dena. T klusterreko operadorea hainbat operadore konektatuk osatzen dute, beren osagaiei dagokienez kanporatuak izan daitezkeenak, eta Φ_i^a bakarra, Φ_{ij}^{ab} bikoitza, Φ_{ijk}^{abc} hirukoitza eta abar sartzen dituzte uhin-funtzioan.

Schrödingerren ekuazioa kluster akoplatuen uhin-funtzioak erabiliz idatz daiteke:

$$H |\Psi_0\rangle = He^{T} |\Phi_0\rangle = Ee^{T} |\Phi_0\rangle \tag{8.15}$$

Teknika "projectivo" bat erabiliz, ekuazio hau Φ_0 erreferentziarekin biderkatu daiteke energia adierazteko.

$$\left\langle \Phi_0 \left| \hat{H} e^{T} \right| \Phi_0 \right\rangle = E \left\langle \Phi_0 \left| e^{T} \right| \Phi_0 \right\rangle = E \tag{8.16}$$

Gehien erabiltzen den kluster akoplatu metodo bat CCSD(T) da, non kitzikapen bakoiti eta bikoitzen tratamendu osoa dagoen, eta hirukoitzen ekarpena, berriz, ez iteratiboki kalkulatzen den, hainbatgorputzaren perturbazioaren teoria erabiliz.

8.4.1 *Domainetan Oinarritutako Tokiko Orbital Parea (DLPNO)*

Kimika konputazionalean, CC metodoak erabiltzeko oztopo bakarra kalkuluek duten kostu konputazional handia da. Diskoaren espazio beharrak molekularen atomo kopuruen laugarren potentzia bezala eskatzen du, eta konputagailuen CPU kalkulu denborak $O(N^6)$ eta $O(N^7)$ eskalatzea dute CCSD eta CCSD(T) kalkuluetarako, hurrenez hurren. Urteetan zehar CC ekuazioetara hurbilketa asko iradoki dira eta modu batera edo bestera gehienek lokalizazio teknikak erabiltzen dituzte [8, 9, 50, 206, 238]. Hurbilketa guztien artean, CC metodo lokalen aldaera batek funtzionatu du. Ikuspegi hori Pulayk eta bere lankideek garatu zuten, tokiko korrelazio-domeinuen kontzeptuan oinarrituta [39, 190, 208, 209], eta Vernerrek eta lankideek garatu eta aplikatu dute [95, 193, 217]. Metodo hauetan, barneko orbital lokalizatuak eta orbital atomiko proiektatuak (PAO) erabiltzen dira espazio birtuala zabaltzeko. Programazio egokiarekin eta atalase zabalarekin, metodo honek eskalatze-lineala aplikatu dezake baliabide konputazional guztiekiko (disko-espazioa, memoria nagusia eta CPU). Boughtonek eta Pulayk sartutako korrelazio-domeinuen definizio estandarrarekin [39], korrelazio-hurbilketa lokalek korrelazio-energiaren 99,6% berreskuratzen dute batez beste [259].

DLPNO-CCSD(T) metodoan korrelazio lokalaren arazoa beste estrategia bat jarraitzen du. Barne espazioa orbital molekular lokalizatuek osatzen duten bitartean, kanpoko espazio korrelatiboa bi orbital naturalen bidez (PNO) hartzen du [168]. PNOak, elektroi bikote bakoitzarentzat ezberdinak diren orbital multzo oso trinko bat dira. PNO metodoan, elektroi bikote bakoitzarentzako orbital naturalak eraik-

itzen dira korrelazio dinamikoaren energiaren hedapen konbergente azkar lortzeko.

8.5 ENERGIAREN DESKONPOSIZIO ANALISIA (EDA)

EDA delakoak AB molekulen A–B interakzio intrintsekoak ditu ardatz, lotura disoziazioaren energiatik (BDE) bereizi behar direnak. Azken termino honek A-B molekularen A eta B fragmentuen erlaxazio geometriko eta elektronikoa dakar oreka geometrian eta oinarrizko egoera elektronikoan. Horrela, EDA molekularen lotura-egoeraren irudikapen fidela ematen du. Energia terminoen batura EDA eta erlaxazio energiak BDE ematen dutenez, behagarri batera konektatuta daude, eta horren ondorioz, EDA tresna erakargarri eta indartsua da fenomeno kimikoak interpretatzeko eta azaltzeko, hala nola egitura molekularra eta erreaktibitatea. Printzipioz, edozein hurbilketa kimiko kuantikorekin batera erabil daiteke. Hemen, DFTrako Kohn-Sham ikuspegian jasotako orbital molekularren eredu kuantitatiboarekin konbinatzen da, korrelazio energia barne hartzen duena. *Ab initio* metodoetarako, EDA Hartree-Fock kalkuluetarako bakarrik garatu da gaur arte. Horregatik erabili ohi da EDA metodoa DFT kalkuluekin batera, hala ere, badira EDAk egiteko beste modu batzuk ere, adibidez, ORCAk DLPNO-CCSD(T) metodoaren bidez garatutakoa.

Lotura-analisiaren EDA ikuspegian A–B lotura baten interakzio energia (ΔE_{int}) berehalako elkarrekintza-energia da, A eta B zatien artean, erreferentzia elektronikoko egoeran eta ABren geometria izoztuan. Elkarrerageneko energia hiru osagai nagusitan banatzen da:

$$\Delta E_{int} = \Delta E_{elstat} + \Delta E_{Pauli} + \Delta E_{orb} \tag{8.17}$$

ΔE_{elstat} terminoa atomoen karga-banaketa nahasigabeen arteko interakzio elektrostatikoari dagokio, normalean erakargarria dena.

Pauli errepultsioa (ΔE_{Pauli}) atomo isolatuen $\rho E(\alpha) + \rho E(\alpha)$ dentsitate elektroien gainposiziotik eratorritako uhin-funtzioaren transformazioari lotutako aldaketa energetikoa bezala sortzen da. Interakzio orbitalak (ΔE_{orb}) kargaren transferentzia deskribatzen du (emaileonartzailea edo orbital okupatuen arteko interakzioak okupatu gabeko orbitalekin, HOMO-LUMO interakzioak barne, besteak beste), polarizazioa (zati baten nahasketa orbital hutsa/okupatua, beste zati baten presentziaren ondorioz) eta bikote elektronikoen loturak (lotura eraketaren ondoriozko egonkortzea) adierazten ditu.

Beste interakzio atomiko mota bat ere badago, sakabanatze-elkarrekintza (ΔE_{disp}), espezieak elkarreragiten dituzten dipolo induzituen arteko indar erakargarrietatik datorrena. Bikote baten arteko sakabanatze-elkarrekintza ahula da, horregatik, lotura kimikorako arbuiagarria kontsideratzen da askotan. Hala ere, espezie elkarreragileak oso handiak badira, elkarrekintza kopurua azkar hazten da eta lotutako indar erakargarriek lotura kimikoaren zati handi bat eman dezakete. Sakabanaketaren interakziorako zuzenketa-termino esplizitu bat erabiltzen bada, Grimmek iradokitako metodoetan bezala [82, 83], EDAren emaitzak ez dira aldatzen eta sakabanaketaren zuzenketa ΔE_{disp} termino gehigarri gisa agertzen da. Sakabanaketaren elkarrekintza funtzionalaren parte bada, EDAren emaitzak aldatuko dira.

8.6 ATOMOEN TEORIA KUANTIKOA MOLEKULETAN (QTAIM)

Atomoei eta talde funtzionalei egotzitako propietate batzuk molekula batetik bestera transferitu daitezkeela ohartzeak funtsezko zeregina izan du kimikaren garapenean. Behaketa horrek oinarri bat ematen die talde funtzionalen gehigarritasun-eskemei, eta taldeak propietate termodinamiko eta espektroskopikoei egiten dien ekarpenaren iraunkortasunaren adibide da. QTAIM [15], Richard F. W. Bader irakasleak eta bere lankideek garatua, behagarri kuantikoetan oinarritzen da, hala nola, elektroi dentsitatea ρ, eta dentsitate energetikoak.

8.6.1 *Dentsitate elektronikoaren topologia*

Dentsitate elektroiaren topologian, nukleoen indar erakargarriak nagusitzen dira, bere ezaugarri topologiko nagusiarekin, maximo garrantzitsu bat nukleo bakoitzaren kokapenean agertuz.

Elektroi-dentsitateko puntu kritikoa (CP) espazioko puntu bat da, non dentsitatearen lehen deribatuak desagertzen diren, hau da:

$$\nabla\rho = i\frac{d\rho}{dx} + j\frac{d\rho}{dy} + k\frac{d\rho}{dz} \rightarrow \begin{cases} = \overrightarrow{0} & (\text{Puntu kritikoan eta } \infty) \\ \neq \overrightarrow{0} & (\text{Gainontzeko puntuetan}) \end{cases} \tag{8.18}$$

non zero bektoreak operadore gradientearen deribatu bakoitza esan nahi duen, ∇, zero dela eta ez bakarrik bere batura. Funtzio eskalar baten gradientea, $\rho(r)$ (8.18 ekuazioa) espazioaren puntu batean bektore bat dago, $\rho(r)$ hazkunde-erritmo handiena duen norabidea adierazten duena, eta norabide horretako hazkunde-erritmoaren adinako magnitudea duena. Nukleo baten posizioan dagoen maximoak CP mota bat osatzen du, hau da, NCP bat.

8.6.2 *Puntu kritikoen sailkapena*

CPa *maila* (ω) eta *sinadura* (σ) arabera sailkatzen da, (ω,σ) eran adierazten dena. Sailkatze maila ρ dituen kurbaduren kopurua da CP puntuan. CP bat, $\omega < 3$ duena, matematikoki ezegonkorra da eta desagertu egingo da mugimendu nuklearrak eragindako dentsitatearen perturbazio txiki baten ondorioz. Horrelako CP bat egoteak (hiru baino maila txikiagokoa) dentsitatearen topologiaren aldaketa adierazten du eta, beraz, egitura molekularraren aldaketa. Horregatik, CP bat $\omega < 3$ balioa duena normalean ez dago orekan eta ia beti $\omega = 3$ aurkitzen da. Lau motatako CP egonkorrak daude, zero ez diren hiru autobalioak dituztenak:

- (3,-3) Hiru kurba negatibo: ρ tokiko maximoa da.
- (3,-1) Bi kurba negatibo: ρ maximoa da dagokion autobektoreak zehaztutako planoan, baina minimoa da hirugarren ardatzean zehar, plano honekiko perpendikularra dena.
- (3,1) Bi kurba positibo: ρ da minimo bat dagokion autobektoreak zehaztutako planoan eta maximo bat hirugarren ardatzean zehar, plano honekiko perpendikularra dena.
- (3,3) Hiru kurba positibo: ρ tokiko minimoa da.

Deskribatutako CP mota bakoitza egitura kimikoko elementu batekin identifikatzen da: (3,-3) nukleo puntu kritikoa NCP; (3,-1) lotura puntu kritikoa BCP; (3,1) eraztun puntu kritikoa RCP; eta (3,3) kaiola puntu kritikoa CCP. 28. Irudiak molekula kubano baten grafiko molekularra erakusten du (lotura bideak eta CPak). Lotura bide kritikoa elektroi dentsitate maximoko lerro bakar bat da, bi atomo kimikoren nukleoak lotzen dituena.

8.6.3 *Lotura propietateak*

Zero-fluxuko azalera bat $\nabla\rho(r)$ ibilbide multzo zehatz batek definitzen du, non kide guztiak puntu bakar batean amaitzen diren, BCP, non $\nabla\rho(r) = 0$ den. BCP amaitu eta azalera interatomiko bat definitzen duen ibilbide multzoaz gain, BCP puntuan ibilbide pare bat sortzen dira, taldeko kide bakoitzarekin, atomoen nukleo batean amaitzen dena. Azken ibilbide pare horrek zehazten du lotura-bidea [13]. Lotura kimikoko elkarrekintzak elektroiaren propietateen eta BCP dentsitate energetikoen arabera ezaugarritzen eta sailkatzen dira, "loturaren propietate" bezala ezagutzen direnak:

- BCP puntuko elektroi dentsitatea (ρ_b). Lotura kimiko baten indarra, bere lotura-ordena (BO), BCPko elektroi-dentsitatean islatzen da, (ρ_b):

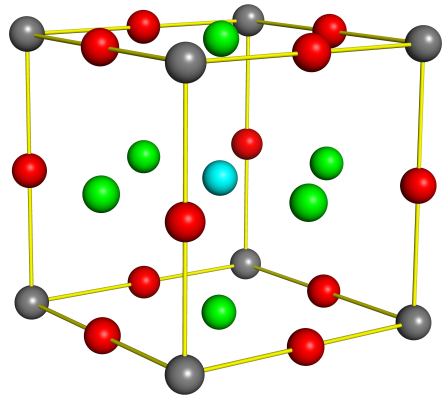

Figure 28: Molekula kubano baten grafiko molekularra, lotura bideak (lerroak) eta CP ezberdinak erakusten dituena: atomo nuklearrak (Karbono atomoak, puntu grisez irudikatuak), lokarriak (puntu gorriak), eraztunak (puntu berdeak) eta kaiola (puntu urdina) CP.

$$BO = [A(\rho_b - B)] \tag{8.19}$$

non A eta B konstanteak diren, loturiko atomoen izaeraren araberako konstanteak diren.

- Loturiko atomo baten erradioa (r_b) eta Lotura Bidearen Luzera. BCP batek A nukleora duen distantziak, A atomoaren "lotura erradioa" zehazten du BCPk definitutako elkarrekintzarekiko, eta $r_{b(A)}$ izendatzen da. Lotura-bidea nukleoen arteko ardatzarekin bat badator, orduan loturiko bi lotura-erradioen baturak, lotura-bidearen luzera deitzen denak, loturaren luzera berdintzen du.

- Elektroi Dentsitatearen laplaziarra BCP nablan. BCPko laplaziera CPko dentsitatearen hiru kurben batura da lotura-bidearekiko perpendikularrak diren biak, $\lambda 1$ eta $\lambda 2$, negatiboak izanik (kon-

bentzioz, $|\lambda_1| > |\lambda_2|$); hirugarrena, berriz, λ_3, lotura-bidean datzana, positiboa da. Kurba negatiboek dentsitatea lotura-bidean zenbateraino kontzentratzen den neurtzen dute, eta kurbadura positiboak atomoen arteko azaleran zenbateraino agortzen den eta banakako oinarri atomikoetan kontzentratzen dena neurtzen du.

8.7 ELKARREKINTZA EZ-KOBALENTE GRAFIKOAK (NCI)

Historikoki, NCI hainbat ekarpen ezberdinetan deskonposatu dira: elektrostatika, polarizazioa, trukatzea eta karga-transferentzia efektuak [126]. Hurbilketa posible bat NCI $\rho(r)$ elektroi dentsitatearen bidez eta bere eremu eskalar eta bektorialen bidez aztertzea da. Ikuspegi horrek bat egiten du DFT kontzeptu giltzarriarekin: elektroiaren dentsitateak lotura zuzena izan dezake energiarekin, eta, beraz, elkarrekintzak dentsitatetik ere aztertzeko gai izan beharko genuke. Kasualitatez, interakzio-energiari egindako ekarpen guztiak, zero ordenako elektrostatikoak izan ezik, $\rho(r)$ deformazio batekin lotu daitezke. Metodologia hauek hainbat interpretazio abantaila natural dituzte. Inbarianteak dira orbital molekularren transformazio unitarioei dagokienez, sendoak izaten dira oro har, eta bisualki azter daitezke, hiru dimentsiokoak direlako.

NCI analisiak indize bat eskaintzen du, elektroiaren dentsitatean eta bere deribatuetan oinarritua, elkarrekintza ez-kobalenteak identifikatzea ahalbidetzen duena [57, 118]. NCI indizea dentsitate lokaleko gradienteko 2D grafiko batean oinarritzen da, s, eta elektroi dentsitatean, ρ, non

$$s = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla\rho|}{\rho^{4/3}} \tag{8.20}$$

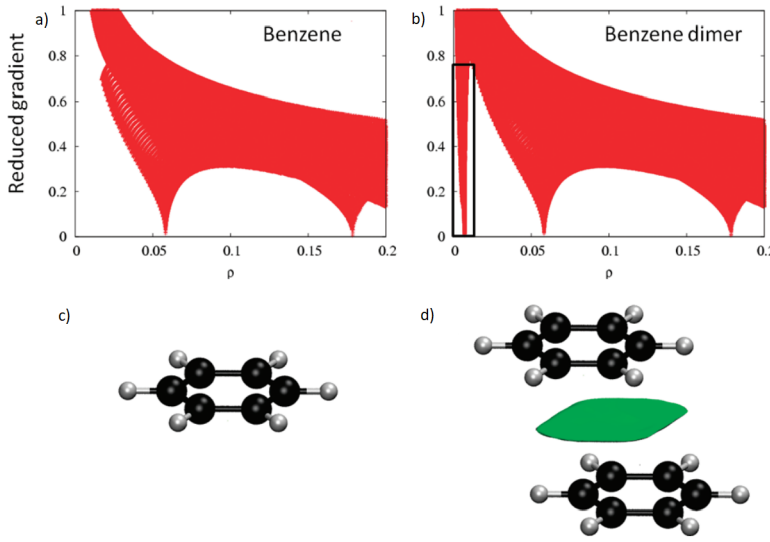

Figure 29: Dentsitate txikiko portaeraren alderaketa Bentzeno a) monomerorako eta b) dimerorakoa; Dentsitate baxuko balioetan partikularitatea agertzen da s-n. c) Bentzeno monomeroa. d) Bentzenoaren dimeroan interakzio molekularreko gainazal baten agerpena, $s(\rho)$) grafikoan dagoen berezitasun gehigarriari lotua. Isoazalera $s = 0.7$ au eta $\rho < 0.01$ a erabiliz sortu da. [57]

Interakzio inter- edo intramolekular ahul bat dagoenean, aldaketa erabakigarria gertatzen da atomo elkarreragileen arteko gradiente murriztuan, CP dentsitatea sortuz elkarrekintza duten zatien artean. Kanalak $s(\rho)$ bakoitzean agertzen dira CP bakoitzari lotuta. Dentsitate baxuetan s-ren portaera ρ-k menderatzen duenez, s-k desbideratzeko joera du, CP dentsitate baten inguruko eskualdeetan izan ezik, non $\nabla\rho$ nagusi den eta s zerora hurbiltzen den. Datu hori 6. irudian nabarmentzen da, eta horrek erakusten du monomero eta dimero baten arteko desberdintasunik handiena dentsitate baxuko kanal malkartsua dela. Espazio errealean ezaugarri hau sortzen duten puntuak bilatzen ditugunean, eskualde ez-kobalentea argi eta garbi agertzen da molekula konplexuetan (isoazalera berdea 29d irudian).

Kanaletako elektroien dentsitatea sakonago aztertu behar da kanal horien jatorria zehazteko (interakzio esterikoak, hidrogeno loturak, etab.). Kanalen barruan dauden elektroien dentsitate-balioak elkarrekintzaren indarraren adierazle dira. Hala ere, elkarreragin erakargarri eta alderakorrak (hau da, lotura hidrogenoa eta errepultsio esterikoa, hurrenz hurren) dentsitate/espazio gradiente murriztuaren eskualde berean agertzen dira. Elkarreragin erakargarri eta aurkakoen artean bereizteko, dentsitatearen bigarren deribatuak aztertzen ditugu aldakuntzaren ardatz nagusian zehar.

Dibergentziaren teorema oinarri hartuta [5], dentsitatearen laplaziar zeinuak ($sign\nabla^2\rho$) adierazten du dentsitatearen fluxu gradiente garbia sartu ($\nabla^2\rho < 0$) edo atera ($\nabla^2\rho > 0$) egiten den bolumen infinitesimal bat uzten du erreferentzia-puntu baten inguruan. Horregatik, $\nabla^2\rho$ zeinuak zehazten du dentsitatea puntu horretan kontzentratzen edo agortzen den, inguruarekin alderatuta. Elkarreragin ahul mota ezberdinak bereizteko, ezin da erabili Laplaziarraren zeinua, nukleoen ekarpen negatiboek menderatzen baitituzte [14]. Horren ordez, bere gehienezko aldakuntza ardatzetan zehar Laplaziarrari egindako ekarpenak aztertu behar dira. Ekarpen horiek dentsitate elektronikoko matrize Hessiarraren (bigarren deribatua) λ_i autobaloreak dira, $\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$, ($\lambda_1 < \lambda_2 < \lambda_3$). Nukleoetan, autobalio guztiak negatiboak dira, eta horietatik urrun $\lambda > 0$. Molekuletan, λ_3 balioak nukleoen arteko norabidean aldatzen dira, eta λ_1 eta λ_2, berriz, planoaren dentsitate aldakuntza normala λ_3 autobaloreetan. Bitxia bada ere, bigarren autobalioa (λ_2) positiboa edo negatiboa izan daiteke, elkarrekintza motaren arabera. Alde batetik, lotura-elkarrekintzek, hidrogeno-loturek esaterako, loturarekiko perpendikular dagoen dentsitate-metaketa dute ezaugarri, eta $\lambda_2 < 0$. Lotura gabeko elkarrekintzak, hala nola errepultsio esterikoak, dentsitate-agortzea eragiten dute, $\lambda > 0$. Azkenik, vdW-eko elkarrekintzek dentsitate-gainjartzearen deuseztatzea dute ezaugarri, λ_0 ematen duena. Horrela, λ_2 zeinuaren analisiak elkarreragin ahulen

mota desberdinak bereizteko aukera ematen digu, dentsitateak berak indarra ebaluatzeko aukera ematen digun bitartean.

Era berean, interakzioak kualitatiboki sailka daitezke NCI karakterizazioa osatzeko. Aurreko 2D irudikapenak $s(r)$ dentsitatearen biderkaduraren eta $H(\rho, r)$ bigarren autobalorearen zeinuaren aurka trazatuz zabal daitezke. Normalean $sign(\lambda_2)\rho(r)$ bezala adierazten da. λ_2-ren zeinuak karga metatzea neurtzen du elkarrekintzaren plano perpendikularrean. $sign(\lambda_2)\rho(r)$ funtzioa interakzio motak sailkatzeko gai da: oso erakargarriak (adibidez, hidrogeno bonuak) < 0, van der Waals $\simeq 0$ (zeinua edozein dela ere) eta talka esterikoak > 0.

9

LABURPENA ETA AZKEN ONDORIOAK

G-quadruplexen interesa areagotu du azken hamarkadetan tumoreen aurkako terapiak ordezkatzeko erabil daitekeelako, zeinek albo-ondorio kaltegarriak eragiten baitituzte giza gorputzari. Gainera, bakterioen aurkako erresistentziari aurre egiteko ere erabil litezke. GQren ikerketak bi kategoria nagusitan banatzen dira: i) DNA katearen portaeraren azterketak GQren DNAren bigarren mailako egitura hartzen duenean, baita egituraren toleste eta zabaltze prozesua ere, eta ii) GQ egiturek molekula txikiekin, molekula organiko hutsekin edo metal-konplexuekin duten elkarrekintzaren azterketa. Tesi honetan bigarren ikasketa motan zentratu gara. GQ molekula txikiekin duen elkarrekintza aztertzean, GQ egitura ahalik eta gehien egonkortzea da helburu ohikoena. Egonkortze hori prozesu biologiko desberdinak eten egiten ditu, hala nola transkripzioa edo erreplikazioa, zelulen apoptosia eraginez, bereziki tumore-zelulei eraginez, duten erreplikazio azkarra dela eta. GQekin lotzen diren molekula txikien ebidentzia biofisiko argia dago, dDNA edo GQ sekuentziaren selektibitate aldakorrarekin. Molekula txiki horiek zelula-mekanismo espezifikoetan dituzten ondorioak (hazkundearen inhibizio-efektuak, transkripzioaren eta erreplikazioaren etenaren edo telomerasaren jardueraren alterazioa...), eta itxaropentsuak dira minbiziaren aurkako tratamendu potentzialetan garatzeko, etorkizuneko hainbat erronkari heldu behar zaio GQ bideratzeko kontzeptuaren sendotasuna berresteko. Bigarren mailako DNA egitura hauek aztertzeak konputazio-azterketetatik onuragarriak izan daitezke, eta horrek ligandoen elkarrekintza izaera ulertzen laguntzen du. Konputazio-metodoek hazkunde nabarmena izan dute

azken hamarkadetan, konputazio-potentzia handiagoa eta algoritmo eraginkorragoak garatuz.

Duela gutxi arte, mekanika molekular klasikoa eta mekanika kuantikoko metodoak ziren sistema biologikoak aztertzeko bi teknika konputazional nagusiak. Lehenak milaka atomoko sistema biologiko handiak erabil ditzake pikosekondoetatik milisegundoetara irauten duten simulazioetarako. Hala ere, ez dute esplizituki deskribatzen molekulen egitura elektronikoa, GQ bezalako egitura biologiko ezhoikoak erreproduzitzeko optimizatuak ez dauden funtzio energetiko enpirikoetan oinarrituz. Bestalde, metodo kuantikoak oso metodo konputazional zehatzak dira, egitura elektronikoa esplizituki deskribatzen dutenak. Hala ere, konputazionalki garestiak dira, eta atomo gutxiko sistemak baino ezin dituzte deskribatzen, gainera denbora laburreko simulazioak bakarrik egin daitezke. Azken hamarkadan, metodo horien arteko arrakala ixten joan gara eskalatze-linealeko metodoei esker, sistema handietarako kalkuluak egiteko aukera ematen baitute, hala nola sistema biologiko gehienetarako, zehaztasuna sakrifikatu gabe eta sistema osoaren egitura elektronikoa deskribatuze. Eskalatze-lineako metodoa metodo konputazional bat besterik ez da, non kalkulu-denbora linealki handitzen den O(N) eskalatze-sistemaren tamainarekin. Metodo horiek DFT eta CCSDrekin batera erabiltzeak aukera ematen du eskala handiko sistemak modu zehatzean modelatzeko arrazoizko denbora konputazionalarekin. Tesi honen helburuetako bat eskalatze-linealeko metodo ezberdinek duten zehaztasuna aztertzea da tamainu handiko DNA sistema ezberdinak molekula txikiekin duten elkarreragina dutenean. Tesi honetan erabili diren DNA egitura ezberdinak hauek izan dira: 1) DNA base-paarea, hidrogeno-loturen eta metaketaren bidez elkarreraginez; 2) dDNAren modelo eberdinak eta phen ligando baten arteko elkarreragina; 3) dDNAren sistema oso bat molekula txiki ezberdinekin elkarreragiten duena, Respinomycin D eta [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal konplexuaren 2 isomerrekin; 4) GQren

bi G-tetrad sistemen hainbat egitura, metal alkalino ezberdinekin egonkortuak, eta 5) GQ sistema oso bat, [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal konplexuaren 2 isomerorekin eta MMQ-1 kinakridina konposatu organikoarekin. Bi metodo erdi-enpirikoren errendimendua konparatu dugu, non dispertsio-zuzenketak sartzen diren, PM6-DH2 eta PM7; QM/MM metodoaren B3LYP-D3(GD3BJ)/6-31+G(d,p):AMBER eta M11L/6-31+G(d,p):AMBER teoria mailetan; DLPNO-CCSD(T)/def2-SVP; eta LS-DFT metodoa SIESTA softwarearekin LMKLL/DZDP teoria maila erabiliz.

PM6-DH2 metodoak sistema horien geometriak zuzen deskribatzen dituela frogatu dugu. Aitzitik, PM7 Hamiltondarrak emaitza okerragoak ematen ditu, eta gainera, GQ egiturak ez ditu zehaztasunez erreproduzitzen. Hala ere, bi metodo erdi-enpirikoek zailtasunak dituzte sistema horien elkarreragineko energiak deskribatzeko, metodo zehatzagoen bidez lortutako emaitzei kontraesanak emanez. Bestalde, frogatu dugu SIESTAk duen LMKLL funtzionalaren bitartez (LS-DFT) modu fidagarrian aurreikusten direla DNA sistema ezberdinen geometria eta elkarrekintza-energiak arrazoizko konputazio denboran tarte batean. Gainera, lortutako energia guztiak ORCA softwarea erabiliz egindako DLPNO-single-point kalkuluak egin ditzake besterik ez, baina ehundaka atomorentzat CCSD(T) teoriaren mailan, eta horien emaitzak erreferentziazko datutzat har daitezke. Gure lanak erakusten du LS-DFT metodoak egokiak direla DNA sistema handien azterketa konputazionaletarako. Hau bereziki interesgarria da molekula txikien eta DNAren arteko elkarrekintza deskribatzeko, izan ere, kobalenteak ez diren elkarreragin ahulen, polarizazioaren eta karga transferentziaren deskribapen zehatza behar da sistema eta molekula txikiekin dituen elkarrekintzak ezaugarritzeko. Era berean, algoritmoen garapenaren eremua irekitzen du, eskalatze-lineako teknikak aplikatuz, CCSD(T) mailaren emaitza energetiko zehatzak lortzeko gai direnak, aldez aurretik LS-DFTrekin optimizatutako egituretan.

Erabilitako metodo desberdinak dDNA eta GQ sistemak eta molekula txikiak duten elkarrekintza zehatz-mehatz deskriba zitzaketela frogatu ondoren, GQ eta [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] konplexu metalikoaren bi isomeroen arteko elkarreragina aztertu genuen. Gure ikerketak bi elkarrekintza modu ditu ardatz: amaiera-pilaketa lotura, non konplexua GQren barruan dagoen erabat; eta interkalazio modua. Lortutako emaitza energetikoentzat, formazio energiak kalkulatzeko LS-DFT metodoa erabili genuen eta elkarreragineko energian eta disolbatzaileak duen efektua zehazteko EDA. Ondorioztatu genuen Ax isomeroak nahiago duela interakzioaren amaiera-pilaketa lotzeko modua, Eq isomeroak interkalazioa bultzatzen duen bitartean. Gainera, QTAIM eta NCI ikerketak egin genituen AIMALL softwarearekin, [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] konplexu metalikoaren eta DNAren arteko elkarreragin ahulak ezaugarritzeko. NCI eta QTAIM analisietan oinarrituta, Ax isomeroaren amaiera-pilaketaren alde egin egiten duela berretsi da, elkarreragin ahul gehiago daudelako. Ax isomeroak, tetraten arteko azken metaketaren bidez eta DNA atomoz inguratuta, elkarreragin asko ditu pi-pi. Aldi berean, metalezko multzoaren lokailu osagarri guztiek DNArekiko interakzio garrantzitsuak ezartzen dituzte. Aldiz, Eq isomer-en kasuan, kanonikoa ez den DNAren bigarren mailako egitura kanpotik interakzionatuz, lokailu osagarri guztiek ez dute elkarrekintzan parte hartzen.

Emaitza horiek guztiak konplexu metalikoaren eta GQaren arteko afinitatea bermatzen dute. Gure ikerketa osatzeko, dDNA edo GQ konplexu metalikoaren selektibitatea ere zehaztu nahi genuen, hau da, dDNA eta GQaren arteko afinitate erlatiboa konplexu metaliko mota hauentzat. Horretarako, formazioa, elkarrekintza eta solbatazio energiak alderatzen ditugu eta elkarreragin ahulak aztertzen ditugu bigarren mailako DNAren bi egiturekin. EDArekin lortutako ΔE_{int} emaitzek [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] konplexu metalikoaren eta GQ DNAren arteko elkarrekintza hobea erakusten dute, dDNArekin baino. Sistema guztien arteko elkarreraginik egonkorrena Ax isomeroaren eta GQ

DNAren artekoa da. Eq isomeroak ere GQ DNArekin elkarrekintza-energia negatiboagoa erakusten du dDNArekin baino. Hala ere, Eq isomeroaren kasuan, bigarren mailako bi DNA sistemen arteko alde energetikoa ez da Ax isomer sistemena bezain handia. Solbatazio efektuak dituzten emaitzek konplexu metalikoak GQ DNArekin selektibitate handiagoa izango lukeen ideia indartzen dute, lortutako ΔE_{aq} dDNArena baino 2/3 aldiz handiagoa baita. Gainera, interakzio ahulen analisiak, QTAIM eta NCI erabiliz, froga gehiago gehitzen dizkio hipotesi honi: konplexu metalikoaren interakzioa mesedegarriagoa litzateke GQ DNArekin. Interakzio ahulen analisiak agerian uzten ditu Ax isomeroak GQ DNArekin duen elkarrekintzak, erabat GQ kabitatearen barruan eta DNA atomoz guztiz inguratuta, $\pi - \pi$ elkarrekintzarik zabalena sortzen duela, phen eta tetraden arteko azalera ahalik eta gehien aprobetxatuz. Gainera, lotura osagarriak ere badaude, tetradekin interakzio garrantzitsuak sortuz, ez bakarrik gertukoekin, baita bigarren geruzako tetradekin ere. Gainera, Eq isomeroak GQ DNAtik kanpo eragiten badu, $\pi - \pi$ azalera handia erakutsi du. Nahiz eta lokailu osagarriak ez izan Ax isomer-en kasuan bezain aktiboak, berebiziko garrantzia dute elkarrekintzan. Bestalde, konplexu metalikoak dDNArekin duen elkarreragina ez da GQerako bezain mesedegarria. Ikusi da Ax isomerorako $\pi - \pi$ azalera zabala dagoela, baina soilik base-pareen artean, eta, beraz, ez da GQ DNAren kasuan bezain zabala. Gainera, bakarrik phen lokailua DNA baseen artea dago tartekatua, gainerako lokailuak kanpoaldera orientatzen dira, lotura guztiak ez dute elkarreraginean parte artzen. Eq isomeroari behatutako elkarrekintza ahulek ez dute dDNArako selektibitatea sustatzen. $\pi - \pi$ azalera ez da oso zabala, konplexu metalikoa ez dagoelako erabat tartekatuta DNA baseen artean.

Hortaz aparte, lokailuetako batzuek ez dute elkarrekintzan parte hartu, eta aztertutako sistema guztien arteko elkarrekintza modurik itxaropen gutxiena aurkezten duena izan da Hori guztia kontuan hartuta, ondorioztatzen dugu [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] konplexu

metalikoak DNAren bigarren mailako bi egiturekin afinitatea izango duela eta biak egonkortuko dituela. Hala ere, GQ ez-kanonikoa den DNAren bigarren mailako egiturarekin selektibitatea izango du, GQ DNArentzat dDNArentzat baino kidetasun handiagoa baitu.

APPENDIX AND BIBLIOGRAPHY

A

PSML PSEUDOPOTENTIALS

In this Appendix the figures and tables obtained from the calculations carried out with the PSML pseudopotentials are compiled, which complement the obtained results in Chapter 3.

A.1 DUPLEX DNA BASE PAIRS

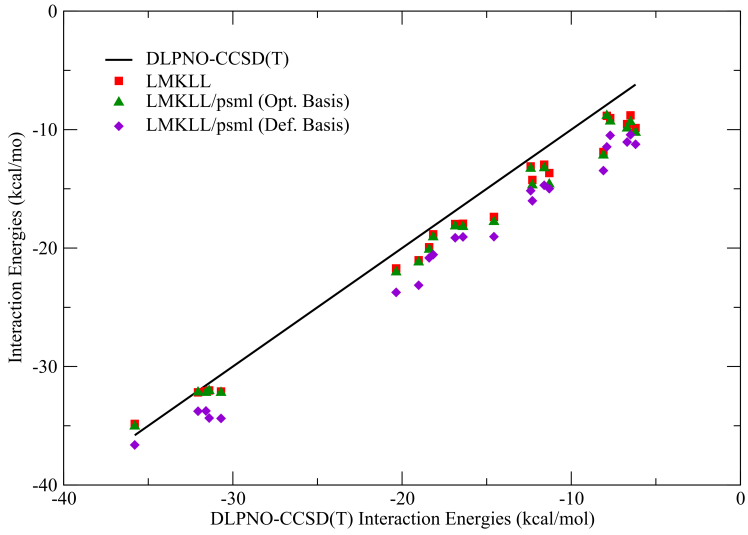

Figure 30: Interaction energies (kcal mol^{-1}) for the DNA base pair benchmark data set structures [120] with psml pseudopotentials. The LMKLL label correspond to the DFT calculations with the optimized pseudopotential and basis set. In the case of LMKLL/psml (Opt. Basis) the psml pseudopotential were used with the optimized basis sets and, in the case of LMKLL/psml (Def. Basis) default basis were used. r^2 value for the LMKLL, LMKLL/psml (Opt. Basis), and LMKLL/psml (Def. Basis) is 0.997, 0.996, and 0.995, respectively.

Table 11: Interaction energies (kcal mol^{-1}) for the different studied DNA base pairs. Abbreviations used in the first column: A, T, C, G – adenine, thymine, cytosine, guanine; m – methyl-; WC, HB – Watson-Crick, Hoogsteen: OG, EG – optimized geometry, experimental geometry.

	System	Reference	LMKLL	LMKLL/psml	
				Opt. Basis	Def. Basis
H-bonded base pairs	G-C WC (OG)	-32.06	-32.18	-32.18	-33.77
	mG-mC WC (OG)	-31.59	-32.09	-32.21	-33.75
	A-T WC (OG)	-16.86	-17.97	-18.16	-19.13
	mA-mT WC (OG)	-18.16	-18.84	-19.08	-20.56
	A-T WC (EG)	-16.40	-17.95	-18.21	-19.06
	G-C WC * (EG)	-35.80	-34.84	-35.05	-36.62
	A-T WC (EG)	-18.40	-19.93	-20.13	-20.83
	G-A HB (EG)	-11.30	-13.67	-14.62	-14.99
	C-G WC (EG)	-30.70	-32.11	-32.21	-34.38
	G-C WC (EG)	-31.40	-32.02	-32.05	-34.35
	MAE	–	1.08	1.27	2.48
Stacked base pairs	G-C (OG)	-19.02	-21.03	-21.20	-23.14
	mG-mC (OG)	-20.35	-21.72	-22.04	-23.74
	A-T (OG)	-12.30	-14.25	-14.70	-16.01
	mA-mT (OG)	-14.57	-17.37	-17.79	-19.04
	A-T (EG)	-8.10	-11.89	-12.18	-13.46
	G-C (EG)	-7.90	-8.84	-8.83	-11.46
	A-C (EG)	-6.70	-9.54	-9.90	-11.05
	T-G (EG)	-6.20	-9.87	-10.26	-11.24
	C-G (EG)	-7.70	-9.03	-9.31	-10.49
	A-G (EG)	-6.50	-8.80	-9.36	-10.44
	C-G (EG)	-12.40	-13.11	-13.30	-15.16
	G-C (EG)	-11.60	-12.96	-13.25	-14.70
	MAE	–	2.09	2.40	3.88

*The geometries of both GC WC (EG) pairs are identical.

A.2 STACKED BASE PAIRS WITH THE INTERCALATED PHEN LIGAND

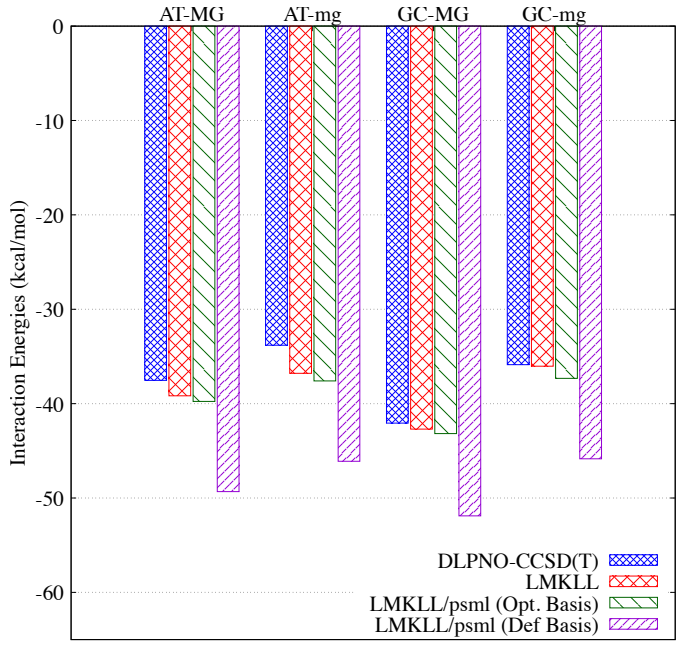

Figure 31: Interaction energies (kcal mol^{-1}) for the phen/DNA system with psml pseudopotentials. The LMKLL label corresponds to the LS-DFT calculations with the optimized pseudopotential and basis set. In the case of LMKLL/psml (Opt. Basis) the psml pseudopotential were used with the optimized basis sets and, in the case of LMKLL/psml (Def. Basis) default basis were used.

Table 12: Interaction energies (kcal mol^{-1}) of the stacked base pairs with the intercalated phen ligand. A-T/phen/T-A MG and A-T/phen/T-A mg corresponds to Adenine-Thymine base pair system with intercalated phen via the Major groove (MG) and minor groove (mg), while G-C/phen/C-G MG and G-C/phen/C-G mg corresponds to Guanine-Cytosine base pair system with phen intercalated through the Major groove (MG) and minor groove (mg). The third column corresponds to the LS-DFT calculations with the optimized pseudopotential and basis set. For the fourth and fifth column psml pseudopotential have been used but, in the fourth column optimized basis sets were used and for the fifth the default basis were used. The Mean Absolute Error (MAE) was calculated taking the DLPNO-CCSD(T) energies as reference values.

System	Reference	LMKLL	LMKLL/psml	
			Opt. Basis	Def. Basis
A-T/phen/T-A MG	-37.53	-39.17	-39.76	-49.32
A-T/phen/T-A mg	-33.81	-36.79	-37.59	-46.11
G-C/phen/C-G MG	-42.06	-42.70	-43.20	-51.89
G-C/phen/C-G mg	-35.87	-36.03	-37.33	-45.84
MAE		0.94	2.15	10.97

B

QTAIM AND NCI PLOTS FOR THE INTERACTION BETWEEN MO-BASED METAL COMPLEX AND G-QUADRUPLEX DNA

In this Appendix, the EDA with the B3LYP-D3/DZP, and QTAIM and NCI plot figures and tables for the interaction between the two isomers, equatorial and axial, of the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex and G-quadruplex DNA have been compiled, which complements the obtained results in Chapter 4.

Table 13: Different contributions in the EDA at the B3LYP-D3/DZP level for the most representative Eq and Ax systems studied in chapter 4. $\Delta E_{aq} = \Delta E_{int} + \Delta E_{Solv}$ and $\Delta E_{Solv} = E_{totalsystem} - E_{solv(GQ)} - E_{solv(metalcomplex)}$. Energy contributions are given in kcal mol^{-1}.

System	Eq1	Eq2	Eq3	Eq4	Ax1	Ax2	Ax3	Ax4
ΔE_{elstat}	-96.2	-88.8	-57.3	-93.2	-116.0	-114.5	-116.6	-80.3
ΔE_{Orb}	-52.7	-49.2	-42.1	-47.4	-66.2	-58.5	-66.8	-49.4
ΔE_{Disp}	-74.5	-71.1	-69.5	-74.9	-106.7	-106.0	-101.3	-90.7
ΔE_{Pauli}	107.3	101.3	95.1	103.8	174.3	176.5	180.5	112.7
ΔE_{int}	-116.1	-107.8	-73.9	-111.7	-114.5	-102.5	-104.3	-107.7
E_{Solv} (DNA)	-493.9	-508.2	-511.0	-525.8	-464.7	-433.4	-467.6	-479.4
E_{Solv} (Complex)	-29.4	-29.4	-22.7	-28.5	-25.8	-23.4	-24.0	-23.0
E_{Solv} (Total)	-468.3	-481.0	-499.2	-505.7	-462.1	-430.7	-459.6	-476.3
ΔE_{Solv}	55.0	56.7	34.5	48.6	28.4	26.2	32.0	26.1
ΔE_{aq}	-61.1	-51.1	-39.4	-63.1	-86.1	-76.4	-72.3	-81.6

B.1 QTAIM

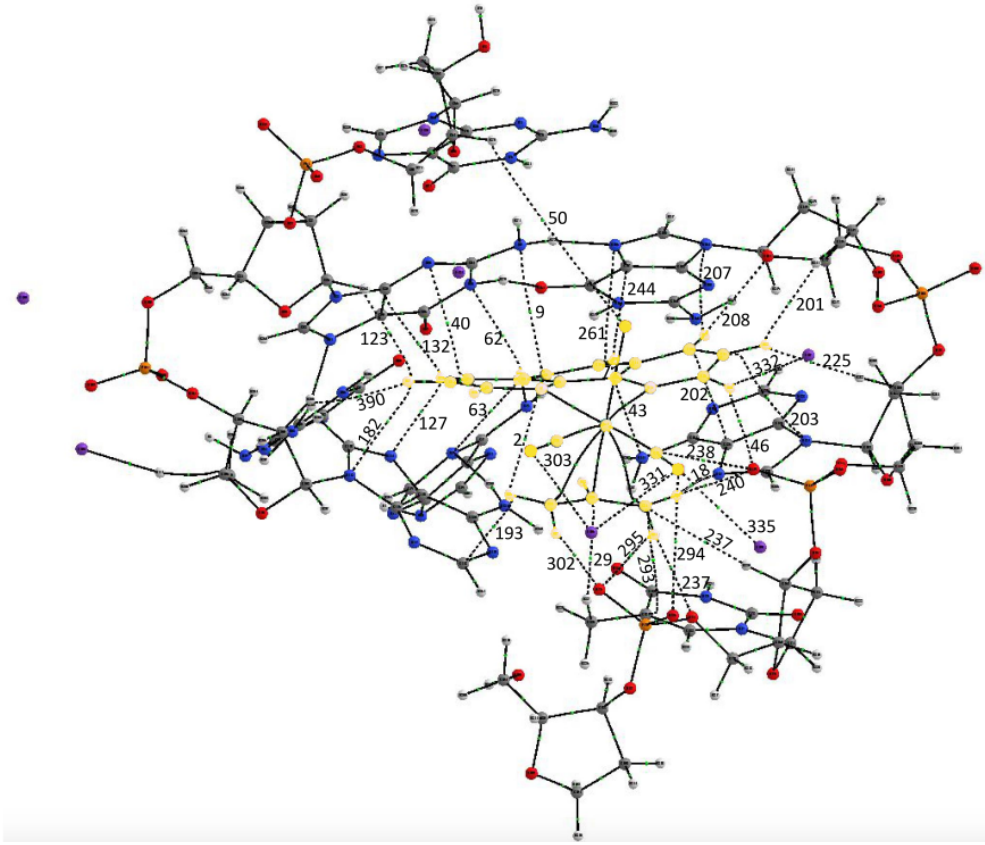

Figure 32: QTAIM topology for the reduced model corresponding to the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] Eq1 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 14: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Eq isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Eq1.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-H_{sugar}	50	0.00004	0.00011	3.11
O_{CO}-N_{base}	18	0.00316	0.01050	3.10
O_{CO}-K	335	0.01160	0.05484	2.84
O_{CO}-K	303	0.01150	0.05221	2.86
O_{CO}-K	331	0.00754	0.03399	3.06
O_{CO}-$O_{phosphate}$	294	0.00287	0.00978	3.74
C_{CO}-$O_{phosphate}$	238	0.00947	0.03669	2.90
C_{phen}-C_{base}	46	0.01128	0.03791	2.98
C_{phen}-N_{base}	2	0.00513	0.01565	3.44
C_{phen}-N_{base}	203	0.00650	0.02187	3.35
C_{phen}-N_{base}	9	0.00622	0.01940	3.38
C_{phen}-N_{base}	40	0.00858	0.02831	3.15
C_{phen}-N_{base}	261	0.00763	0.02567	3.20
H_{phen}-C_{base}	244	0.00791	0.02870	2.69
H_{phen}-C_{base}	132	0.00905	0.03157	2.66
H_{phen}-N_{base}	62	0.00953	0.03169	2.69
H_{phen}-N_{base}	207	0.00761	0.02647	2.70
H_{phen}-N_{base}	182	0.00616	0.02100	2.88
H_{phen}-N_{base}	127	0.01369	0.05015	2.38
H_{phen}-N_{base}	63	0.00375	0.01156	3.10
H_{phen}-N_{base}	43	0.01033	0.03649	2.66
H_{phen}-H_{base}	71	0.00586	0.02341	2.22
H_{phen}-H_{sugar}	390	0.00285	0.00954	2.64
H_{phen}-H_{sugar}	123	0.00406	0.01342	2.56
H_{phen}-H_{sugar}	201	0.00422	0.01457	2.53
H_{phen}-H_{sugar}	225	0.00352	0.01090	2.58
H_{phen}-O_{sugar}	208	0.00988	0.03880	2.55
H_{phen}-K	332	0.00298	0.01324	3.00
H_{phen}-$O_{phosphate}$	202	0.01141	0.03807	2.39
C_{allyl}-C_{base}	29	0.00296	0.00782	3.88
C_{allyl}-H_{sugar}	237	0.00450	0.01633	2.94
H_{allyl}-C_{base}	193	0.00929	0.03503	2.57
H_{allyl}-$O_{phosphate}$	240	0.00882	0.03678	2.62
H_{allyl}-$O_{phosphate}$(s)	293	0.00922	0.03346	2.45
H_{allyl}-$O_{phosphate}$	295	0.00750	0.02733	2.64
H_{allyl}-$O_{phosphate}$	302	0.01680	0.05889	2.13

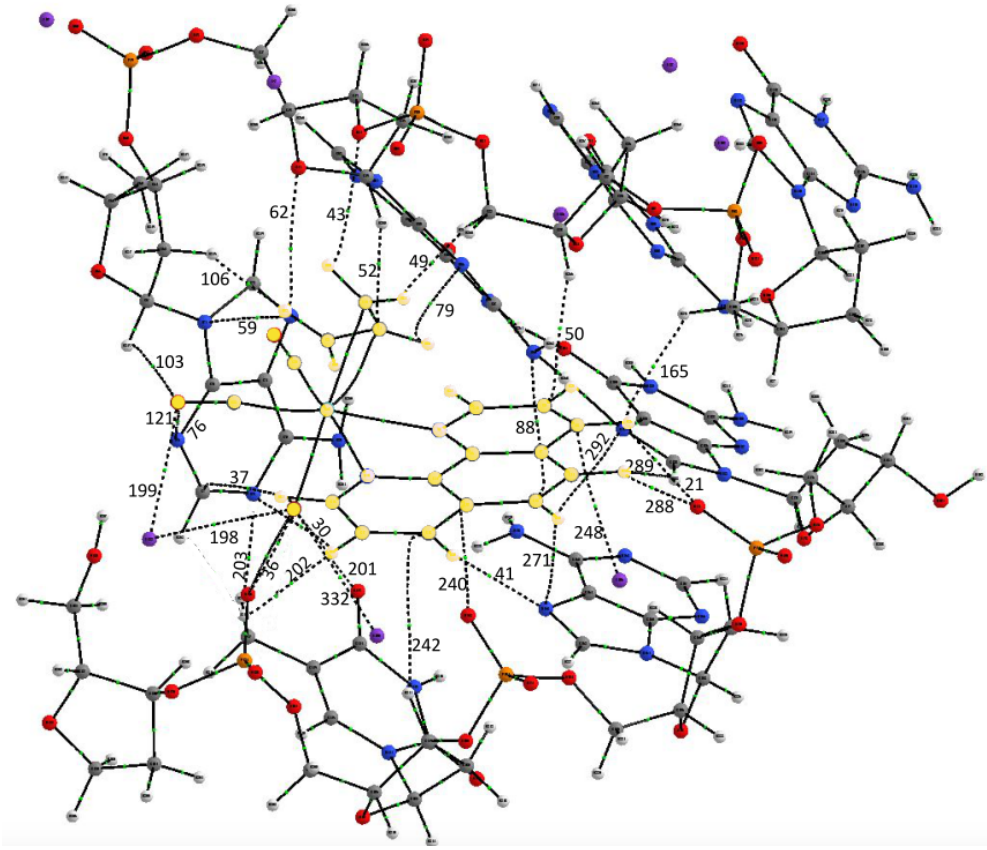

Figure 33: QTAIM topology for the reduced model corresponding to the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] Eq2 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 15: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Eq isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Eq2.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-K	332	0.011567	0.046362	3.21
Br-K	198	0.004633	0.016118	3.70
Br-$O_{phosphate}$	201	0.004199	0.012738	3.80
Br-$O_{phosphate}$	203	0.002673	0.007857	4.13
O_{CO}-K	199	0.009610	0.045253	2.92
O_{CO}-N_{base}	121	0.002667	0.010407	3.63
C_{phen}-N_{base}	88	0.007648	0.022838	3.21
C_{phen}-N_{base}	41	0.013170	0.041336	2.37
C_{phen}-H_{sugar}	50	0.006758	0.022981	2.76
C_{phen}-H_{sugar}	242	0.001438	0.004265	3.64
C_{phen}-$O_{phosphate}$	240	0.009478	0.035755	2.97
C_{phen}-K	248	0.004188	0.016788	3.42
H_{phen}-C_{base}	37	0.005745	0.021103	2.80
H_{phen}-N_{base}	292	0.005211	0.018530	2.88
H_{phen}-N_{base}	271	0.009763	0.031721	2.54
H_{phen}-N_{base}	30	0.013657	0.045906	2.38
H_{phen}-H_{base}	289	0.005269	0.020874	2.30
H_{phen}-H_{base}	202	0.006451	0.026001	2.28
H_{phen}-H_{sugar}	165	0.004065	0.014128	2.52
H_{phen}-H_{sugar}	21	0.009872	0.037136	2.46
H_{phen}-H_{sugar}	288	0.018487	0.069172	2.13
H_{phen}-$O_{phosphate}$	36	0.015420	0.069556	2.28
C_{allyl}-H_{sugar}	52	0.015849	0.052772	2.28
H_{allyl}-N_{base}	79	0.007553	0.026567	2.78
H_{allyl}-N_{base}	59	0.009296	0.033103	2.64
H_{allyl}-O_{sugar}	62	0.003903	0.015562	2.98
H_{allyl}-O_{sugar}	43	0.004580	0.017538	2.90
H_{allyl}-H_{sugar}	106	0.011874	0.038227	1.93
H_{allyl}-H_{sugar}	49	0.015151	0.057778	1.95

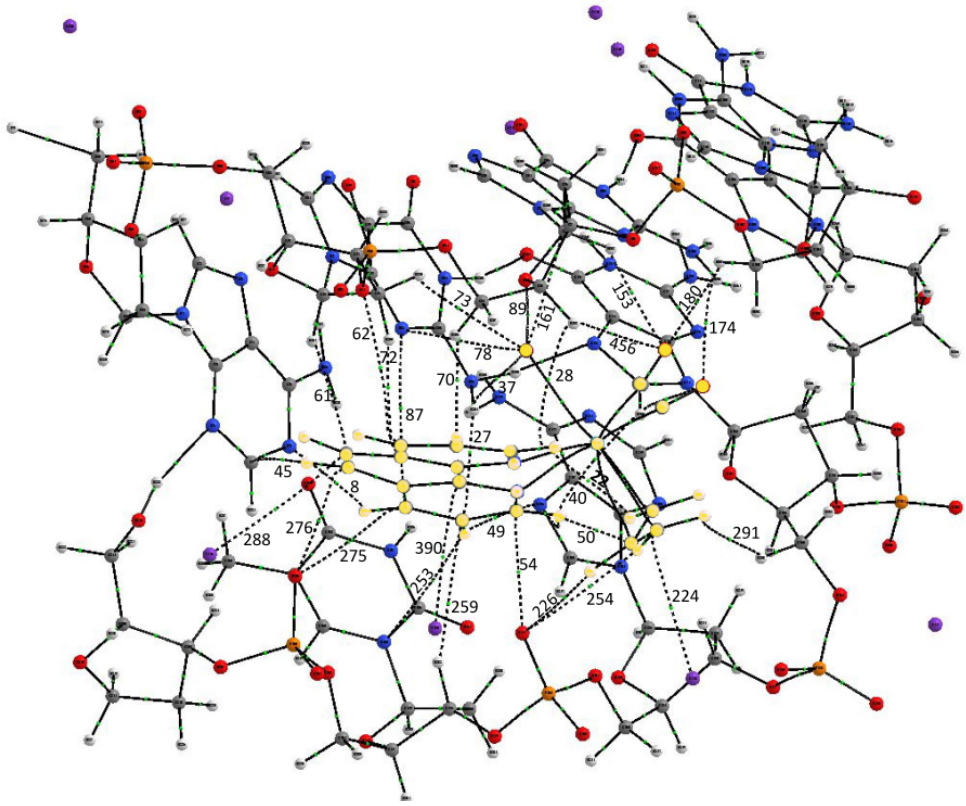

Figure 34: QTAIM topology for the reduced model corresponding to the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] Eq3 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 16: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Eq isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Eq3.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-C_{base}	161	0.0101	0.0367	3.23
Br-N_{base}	78	0.0079	0.0236	3.26
Br-O_{base}	89	0.0065	0.0225	3.43
Br-H_{base}	73	0.0063	0.0205	3.02
Br-H_{base}	37	0.0062	0.0188	3.11
O_{CO}-H_{base}	456	0.0047	0.0191	2.78
O_{CO}-H_{base}	153	0.0075	0.0295	2.53
O_{CO}-H_{base}	174	0.0088	0.0368	2.40
O_{CO}-H_{sugar}	180	0.0053	0.0225	2.65
C_{phen}-K	288	0.0028	0.0103	3.71
C_{phen}-K	390	0.0053	0.0210	3.35
C_{phen}-N_{base}	87	0.0033	0.0101	3.70
C_{phen}-H_{base}	61	0.0065	0.0205	2.85
C_{phen}-H_{base}	27	0.0123	0.0432	2.42
C_{phen}-H_{sugar}	72	0.0091	0.0315	2.58
C_{phen}-H_{sugar}	70	0.0118	0.0411	2.44
C_{phen}-H_{sugar}	259	0.0031	0.0096	3.14
C_{phen}-$O_{phosphate}$	275	0.0083	0.0339	3.07
C_{phen}-$O_{phosphate}$	276	0.0132	0.0545	2.77
C_{phen}-$O_{phosphate}$	54	0.0120	0.0449	2.79
C_{phen}-$O_{phosphate}$	62	0.0035	0.0135	3.50
H_{phen}-C_{base}	45	0.0111	0.0394	2.21
H_{phen}-N_{base}	49	0.0134	0.0475	2.35
H_{phen}-N_{base}	253	0.0017	0.0055	3.51
H_{phen}-N_{base}	8	0.0108	0.0366	2.45
H_{phen}-H_{base}	28	0.0031	0.0097	2.77
H_{phen}-H_{base}	40	0.0018	0.0054	2.92
C_{allyl}-K	224	0.0056	0.0217	3.35
H_{allyl}-H_{base}	50	0.0067	0.0270	2.25
H_{allyl}-H_{base}	22	0.0060	0.0233	2.38
H_{allyl}-H_{sugar}	291	0.0036	0.0123	2.59
H_{allyl}-$O_{phosphate}$	254	0.0032	0.0123	3.09
H_{allyl}-$O_{phosphate}$	226	0.0038	0.0145	2.98

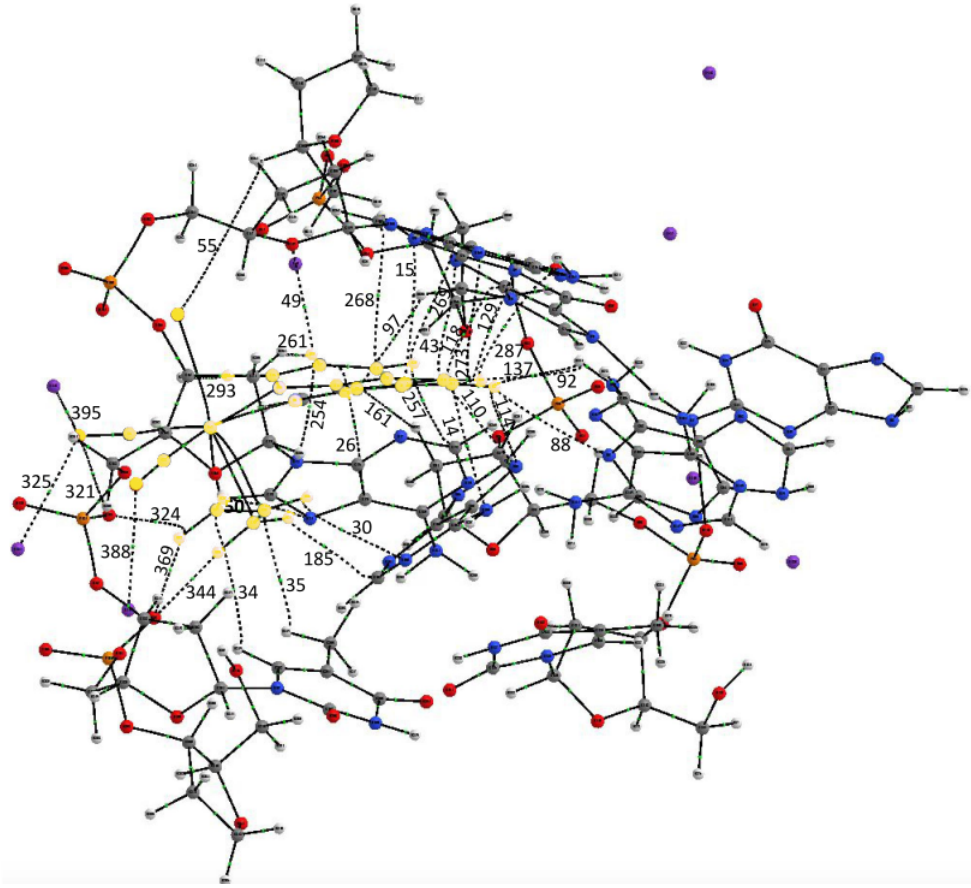

Figure 35: QTAIM topology for the reduced model corresponding to the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ Eq4 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 17: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Eq isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Eq4.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-H_{sugar}	55	0.000098	0.000411	5.09
O_{CO}-K	395	0.011124	0.052422	2.85
O_{CO}-K	325	0.007538	0.037282	3.00
O_{CO}-K	388	0.009575	0.046237	2.90
O_{CO}-$O_{phosphate}$	321	0.009002	0.032667	2.96
C_{phen}-C_{base}	26	0.007879	0.023746	3.17
C_{phen}-C_{base}	268	0.005378	0.017780	3.39
C_{phen}-N_{base}	14	0.010776	0.033383	3.02
C_{phen}-N_{base}	15	0.006097	0.019169	3.38
C_{phen}-N_{base}	129	0.009693	0.031907	3.10
H_{phen}-C_{base}	257	0.004694	0.017463	2.93
H_{phen}-C_{base}	269	0.005695	0.018932	2.90
H_{phen}-C_{base}	273	0.004308	0.013441	3.24
H_{phen}-C_{base}	118	0.009305	0.035578	2.57
H_{phen}-N_{base}	110	0.011633	0.040797	2.54
H_{phen}-N_{base}	88	0.003815	0.012709	3.06
H_{phen}-N_{base}	114	0.006188	0.020909	2.79
H_{phen}-N_{base}	43	0.009394	0.034198	2.71
H_{phen}-N_{base}	254	0.008649	0.029938	2.76
H_{phen}-H_{base}	161	0.002678	0.009073	2.65
H_{phen}-H_{base}	92	0.004294	0.017303	2.42
H_{phen}-H_{base}	137	0.002633	0.010560	2.63
H_{phen}-H_{base}	97	0.002374	0.007742	2.78
H_{phen}-O_{sugar}	49	0.010352	0.039461	2.56
H_{phen}-O_{sugar}	287	0.004089	0.015844	3.05
H_{phen}-H_{sugar}	261	0.010528	0.034937	1.98
H_{phen}-H_{sugar}	293	0.002427	0.007520	2.74
C_{allyl}-H_{base}	34	0.001814	0.005255	3.46
C_{allyl}-H_{base}	35	0.006427	0.022467	2.82
H_{allyl}-C_{base}	185	0.009914	0.036704	2.50
$H_{allyl}N_{base}$	30	0.003568	0.012566	3.16
H_{allyl}-N_{base}	50	0.009145	0.032991	2.56
H_{allyl}-$O_{phosphate}$	324	0.009418	0.041748	2.52
H_{allyl}-$O_{phosphate}$	369	0.011346	0.039481	2.40
H_{allyl}-$O_{phosphate}$	344	0.013726	0.049784	2.26

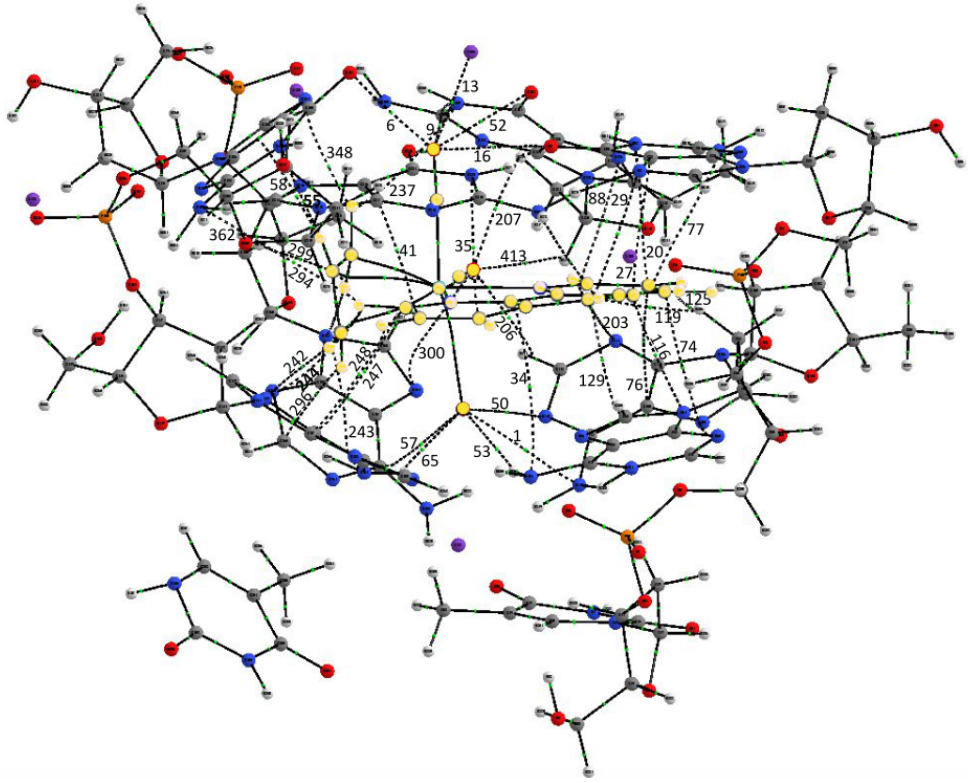

Figure 36: QTAIM topology for the reduced model corresponding to the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ Ax1 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 18: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Ax isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Ax1.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-H_{base}	65	0.019229	0.055594	2.42
Br-H_{base}	53	0.021011	0.060017	2.35
Br-C_{base}	57	0.010058	0.035899	3.23
Br-N_{base}	1	0.011278	0.034918	3.25
Br-N_{base}	50	0.009845	0.030317	3.29
O_{CO}-H_{base}	206	0.008328	0.034541	2.59
O_{CO}-H_{base}	413	0.003141	0.013448	2.95
O_{CO}-C_{base}	207	0.007927	0.031655	3.00
O_{CO}-O_{base}	16	0.002496	0.011639	3.53
O_{CO}-O_{base}	6	0.015365	0.064003	2.62
O_{CO}-N_{base}	300	0.005986	0.021303	3.24
O_{CO}-K	13	0.012688	0.069536	2.72
C_{CO}-O_{base}	52	0.014566	0.054934	2.69
C_{CO}-N_{base}	9	0.009523	0.039315	2.95
C_{phen}-C_{base}	129	0.011621	0.039405	2.97
C_{phen}-C_{base}	76	0.006569	0.020283	3.32
C_{phen}-C_{base}	41	0.006725	0.019220	3.65
C_{phen}-C_{base}	247	0.005277	0.015811	3.45
C_{phen}-N_{base}	116	0.008385	0.027005	3.21
C_{phen}-N_{base}	74	0.003112	0.009790	3.76
C_{phen}-N_{base}	34	0.002286	0.007224	3.94
C_{phen}-N_{base}	77	0.005569	0.016695	3.39
C_{phen}-N_{base}	20	0.006403	0.019293	3.38
C_{phen}-N_{base}	27	0.009215	0.029337	3.12
C_{phen}-N_{base}	29	0.007206	0.022979	3.23
C_{phen}-N_{base}	80	0.006316	0.021673	3.28
C_{phen}-N_{base}	35	0.008322	0.025245	3.19
H_{phen}-C_{base}	248	0.005442	0.019220	2.89
H_{phen}-N_{base}	203	0.004480	0.015689	2.98
H_{phen}-N_{base}	362	0.000697	0.001710	4.14
H_{phen}-N_{base}	244	0.004816	0.017112	2.95
H_{phen}-H_{sugar}	125	0.003801	0.013433	2.64
H_{phen}-H_{sugar}	119	0.004645	0.016872	2.42
H_{allyl}-C_{base}	58	0.007624	0.026125	4.02
H_{allyl}-C_{base}	296	0.010351	0.036429	2.48
H_{allyl}-N_{base}	294	0.009886	0.033516	2.55
H_{allyl}-N_{base}	242	0.009332	0.029290	2.54
H_{allyl}-N_{base}	237	0.002469	0.008853	3.28
H_{allyl}-N_{base}	348	0.009087	0.029675	2.60
H_{allyl}-N_{base}	55	0.007555	0.025071	2.69
H_{allyl}-N_{base}	243	0.006143	0.023050	2.90
H_{allyl}-H_{sugar}	299	0.004024	0.014041	2.58

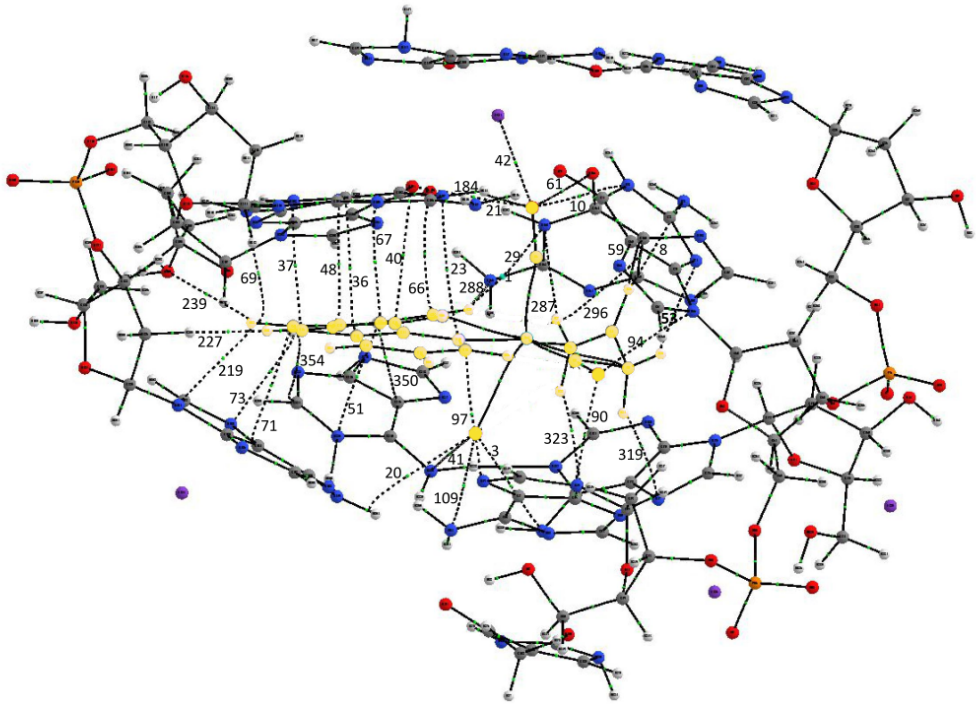

Figure 37: QTAIM topology for the reduced model corresponding to the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ Ax2 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 19: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Ax isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Ax2.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-N_{base}	3	0.011772	0.037597	3.23
Br-N_{base}	41	0.008185	0.023706	3.43
Br-N_{base}	109	0.015482	0.054299	3.06
Br-H_{base}	20	0.006651	0.020038	3.07
O_{CO}-C_{base}	90	0.007437	0.029931	3.02
O_{CO}-N_{base}	94	0.009537	0.034968	2.94
O_{CO}-N_{base}	10	0.016139	0.065246	2.70
O_{CO}-O_{base}	184	0.009302	0.034106	2.92
O_{CO}-O_{base}	61	0.006059	0.024275	3.11
O_{CO}-H_{base}	21	0.007235	0.030220	2.50
O_{CO}-K	42	0.010810	0.056079	2.81
C_{phen}-C_{base}	36	0.007062	0.021188	3.31
C_{phen}-C_{base}	67	0.011837	0.041952	2.91
C_{phen}-C_{base}	71	0.010820	0.035294	3.04
C_{phen}-C_{base}	51	0.008844	0.029928	3.13
C_{phen}-N_{base}	37	0.009494	0.029649	3.09
C_{phen}-N_{base}	48	0.010986	0.034507	3.04
C_{phen}-N_{base}	97	0.006788	0.021565	3.22
N_{phen}-N_{base}	23	0.003826	0.013348	3.58
C_{phen}-O_{base}	40	0.004881	0.017924	3.34
C_{phen}-O_{base}	66	0.009273	0.033335	3.00
H_{phen}-C_{base}	350	0.008139	0.027310	2.71
H_{phen}-C_{base}	354	0.007465	0.030073	2.70
H_{phen}-N_{base}	69	0.007778	0.025595	2.81
H_{phen}-N_{base}	291	0.004755	0.017650	2.91
H_{phen}-N_{base}	288	0.006961	0.021918	2.75
H_{phen}-N_{base}	219	0.008998	0.032869	2.69
H_{phen}-N_{base}	73	0.006365	0.021302	2.82
H_{phen}-H_{sugar}	227	0.014119	0.043001	1.85
H_{phen}-H_{sugar}	239	0.002486	0.007380	2.86
H_{allyl}-C_{base}	296	0.006734	0.023603	2.94
H_{allyl}-C_{base}	59	0.007274	0.025873	2.76
H_{allyl}-C_{base}	323	0.007795	0.028033	2.65
H_{allyl}-N_{base}	287	0.011729	0.038182	2.49
H_{allyl}-N_{base}	8	0.008009	0.023955	2.67
H_{allyl}-N_{base}	53	0.007820	0.025345	2.65
H_{allyl}-N_{base}	319	0.009015	0.029443	2.57

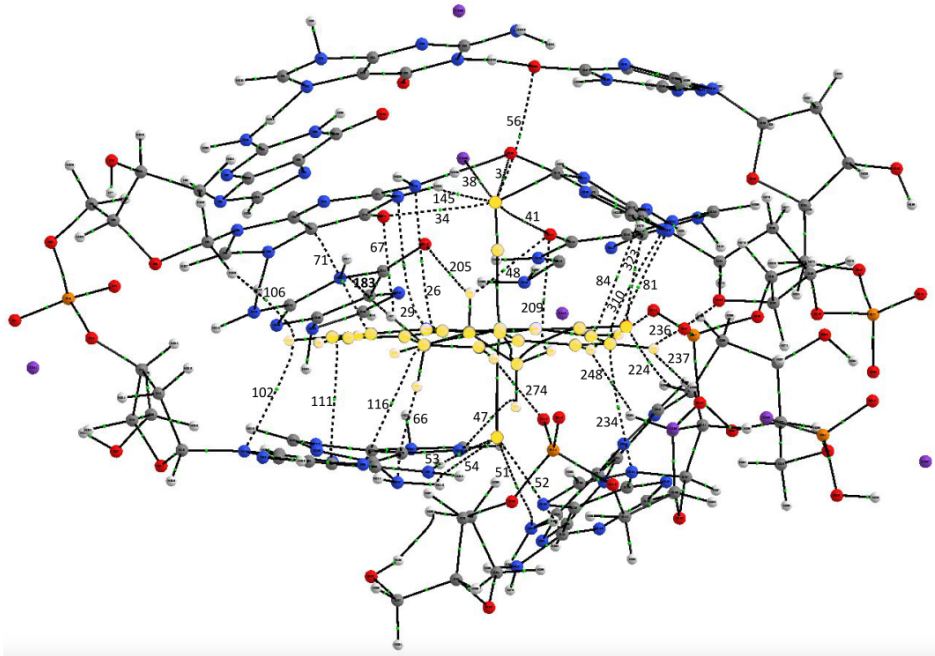

Figure 38: QTAIM topology for the reduced model corresponding to the $[Mo(\eta^3\text{-}C_3H_5)Br(CO)_2(phen)]$ Ax3 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 20: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Ax isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Ax3.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-H_{base}	54	0.015278	0.051962	2.55
Br-H_{base}	53	0.020101	0.055277	2.36
Br-N_{base}	51	0.007338	0.021864	3.54
Br-C_{base}	52	0.013828	0.050743	3.03
O_{CO}-O_{base}	34	0.011697	0.042675	2.85
O_{CO}-O_{base}	3	0.007820	0.029285	3.07
O_{CO}-O_{base}	56	0.002355	0.010525	3.59
O_{CO}-O_{base}	41	0.007592	0.033405	2.96
O_{CO}-H_{base}	145	0.018609	0.065108	2.08
O_{CO}-H_{base}	224	0.009419	0.041482	2.50
O_{CO}-N_{base}	323	0.010636	0.036603	2.91
O_{CO}-K	38	0.015629	0.077961	2.69
C_{phen}-C_{base}	111	0.009290	0.028999	3.14
C_{phen}-C_{base}	71	0.009572	0.030125	3.11
C_{phen}-C_{base}	84	0.007089	0.023861	3.22
C_{phen}-N_{base}	81	0.008397	0.026810	3.14
C_{phen}-N_{base}	26	0.003682	0.011285	3.66
C_{phen}-N_{base}	116	0.006218	0.020112	3.29
C_{phen}-N_{base}	234	0.007036	0.021753	3.26
N_{phen}-N_{base}	29	0.006145	0.019397	3.35
H_{phen}-N_{base}	310	0.007571	0.026117	2.79
H_{phen}-N_{base}	102	0.005682	0.020278	2.93
H_{phen}-H_{base}	237	0.009011	0.033741	2.06
H_{phen}-H_{sugar}	106	0.003801	0.014157	2.64
H_{phen}-H_{sugar}	248	0.005066	0.019033	2.34
H_{phen}-O_{sugar}	236	0.010594	0.038612	2.43
H_{phen}-$O_{phosphate}$	274	0.008358	0.030772	2.48
H_{allyl}-N_{base}	47	0.013221	0.049247	2.47
H_{allyl}-N_{base}	66	0.013313	0.042683	2.40
H_{allyl}-N_{base}	183	0.006946	0.026586	2.78
H_{allyl}-O_{base}	67	0.009803	0.035168	2.47
H_{allyl}-O_{base}	205	0.013081	0.044814	2.33
H_{allyl}-O_{base}	48	0.011233	0.039635	2.43
H_{allyl}-O_{base}	209	0.005390	0.022483	2.85

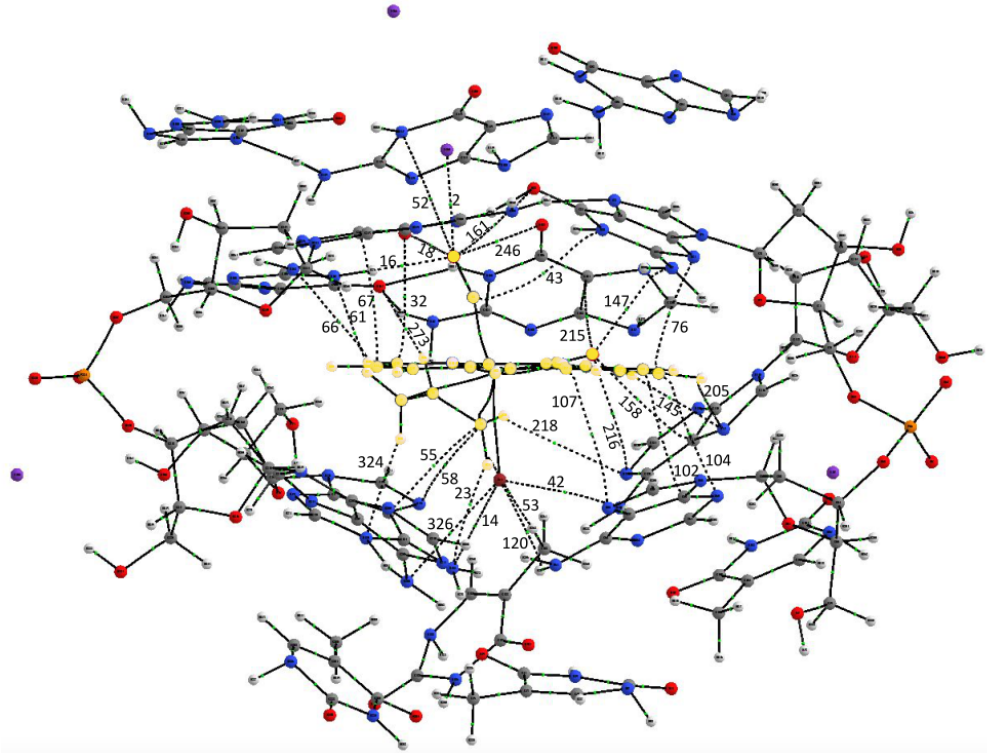

Figure 39: QTAIM topology for the reduced model corresponding to the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] Ax4 system. The metal complex is highlighted in yellow and the Bond Critical Points (BCP) labels are plotted next to the corresponding bond path.

Table 21: Electron Density (ρ), Laplacian ($\nabla_2\rho$) and bond distance (d), in Å, on all BCPs corresponding to the intermolecular weak interactions between the Ax isomer of [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] and GQ DNA model for the system Ax4.

Atom	BCP	ρ	$\nabla_2\rho$	d
Br-N_{base}	14	0.008670	0.024253	3.41
Br-N_{base}	326	0.007063	0.018729	3.53
Br-H_{base}	120	0.011894	0.042669	2.69
Br-N_{base}	42	0.006316	0.017684	3.57
O_{CO}-C_{base}	158	0.006401	0.022974	3.24
O_{CO}-C_{base}	215	0.007631	0.028465	3.84
O_{CO}-N_{base}	147	0.010353	0.041165	2.96
O_{CO}-N_{base}	216	0.006310	0.023987	3.19
O_{CO}-N_{base}	52	0.001014	0.003949	4.13
O_{CO}-O_{base}	161	0.002938	0.012778	3.57
O_{CO}-O_{base}	246	0.009794	0.036841	2.88
O_{CO}-O_{base}	18	0.009039	0.036532	2.90
O_{CO}-H_{base}	16	0.021174	0.067210	1.99
O_{CO}-K	2	0.012917	0.064381	2.77
C_{CO}-N_{base}	43	0.004064	0.011752	3.59
C_{phen}-C_{base}	102	0.006091	0.019819	3.28
C_{phen}-C_{base}	67	0.007546	0.021979	3.26
C_{phen}-N_{base}	107	0.004006	0.012289	3.55
C_{phen}-N_{base}	104	0.005568	0.018115	3.37
C_{phen}-N_{base}	76	0.007363	0.023145	3.23
C_{phen}-O_{base}	32	0.005616	0.021824	3.27
H_{phen}-N_{base}	66	0.003516	0.011628	3.09
H_{phen}-N_{base}	61	0.007497	0.024899	2.69
H_{phen}-N_{base}	145	0.007485	0.026496	2.67
H_{phen}-N_{base}	205	0.006315	0.021968	2.79
C_{alyl}-C_{base}	55	0.008840	0.031707	3.15
C_{allyl}-N_{base}	58	0.008890	0.030967	3.15
H_{allyl}-N_{base}	324	0.009409	0.031309	2.56
H_{allyl}-N_{base}	23	0.007935	0.026604	2.66
H_{allyl}-N_{base}	218	0.002580	0.009014	3.28
H_{allyl}-O_{base}	273	0.011858	0.039299	2.36
H_{allyl}-H_{base}	53	0.003095	0.010196	2.72

B.2 NCI PLOTS

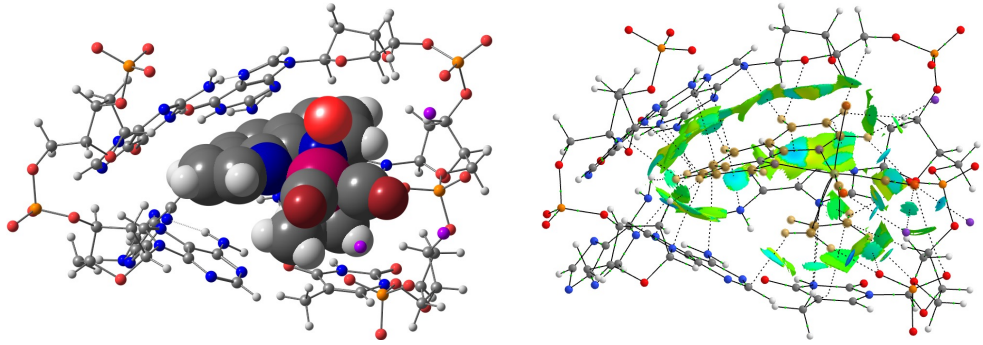

Figure 40: On the left, reduced model of the Eq1 system with the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces (s = 0.5 au) computed for the Eq1.

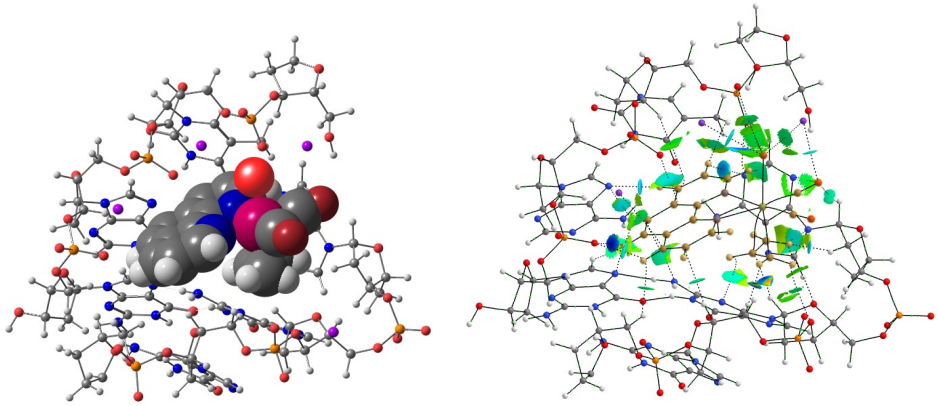

Figure 41: On the left, reduced model of the Eq2 system with the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces (s = 0.5 au) computed for the Eq2.

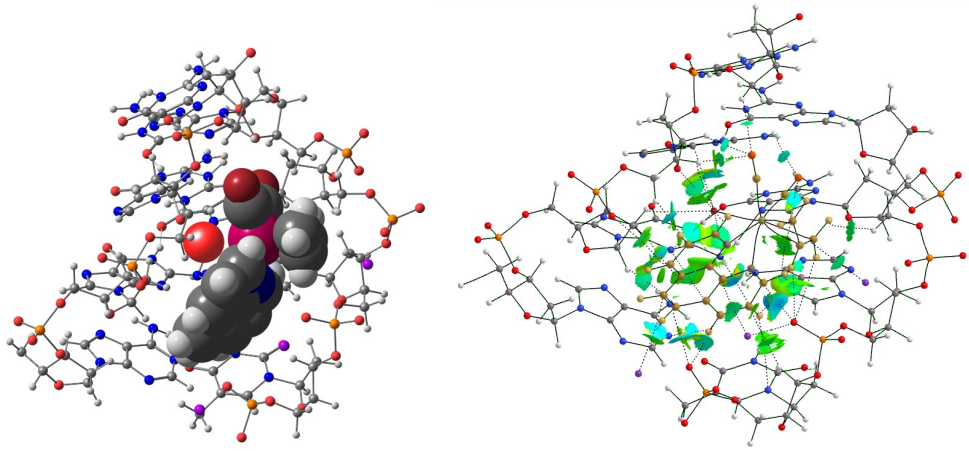

Figure 42: On the left, reduced model of the Eq3 system with the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces (s = 0.5 au) computed for the Eq3.

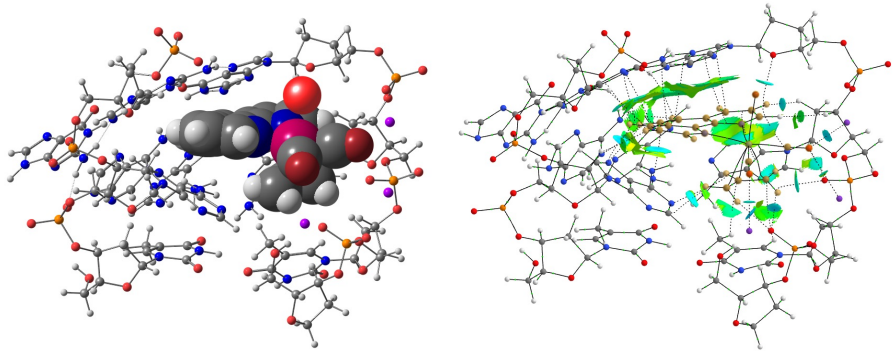

Figure 43: On the left, reduced model of the Eq4 system with the [Mo(η^3-C_3H_5)Br(CO)$_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces (s = 0.5 au) computed for the Eq4.

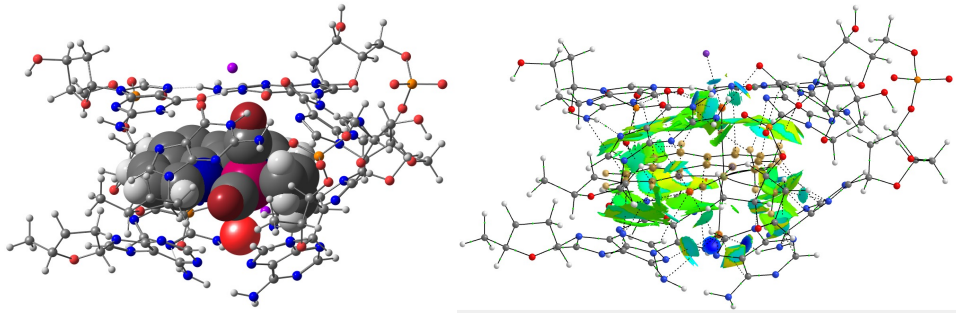

Figure 44: On the left, reduced model of the Ax1 system with the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces (s = 0.5 au) computed for the Ax1.

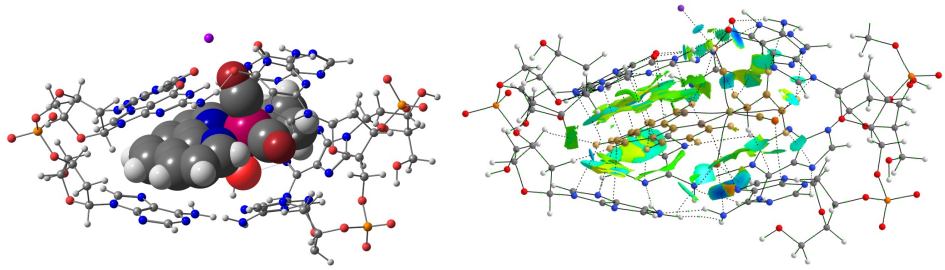

Figure 45: On the left, reduced model of the Ax2 system with the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces (s = 0.5 au) computed for the Ax2.

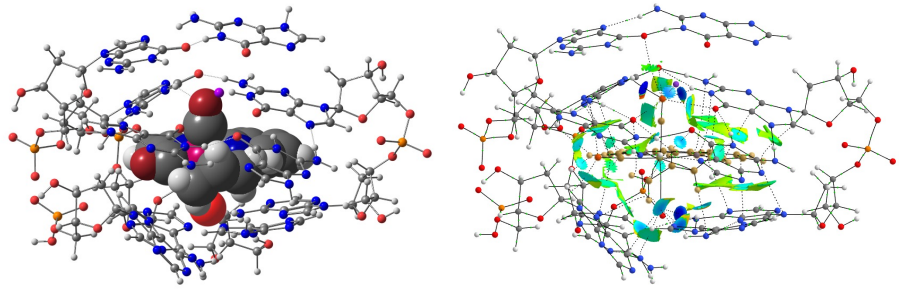

Figure 46: On the left, reduced model of the Ax3 system with the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces (s = 0.5 au) computed for the Ax3.

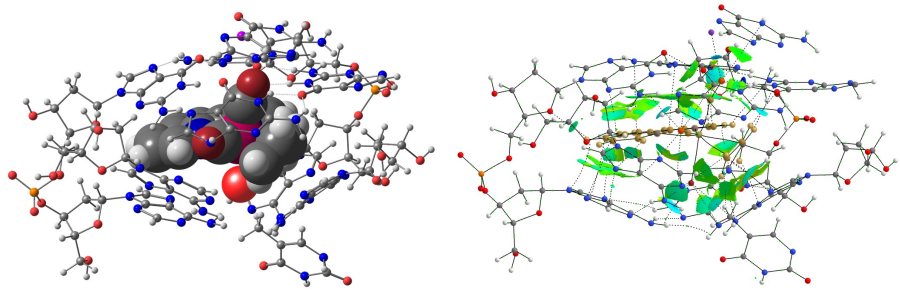

Figure 47: On the left, reduced model of the Ax4 system with the [Mo(η^3-C_3H_5)Br$(CO)_2$(phen)] metal complex represented in spheres. On the right, the NCI index plot with gradient isosurfaces ($s = 0.5$ au) computed for the Ax4.

BIBLIOGRAPHY

[1] ADF2013. *Vrije Universiteit: Amsterdam, The Netherlands*. 2013.

[2] Manali Aggrawal, Hyun Joo, Wanbo Liu, Jerry Tsai, and Liang Xue. "8-Oxo-7,8-dihydrodeoxyadenosine: The first example of a native DNA lesion that stabilizes human telomeric G-quadruplex DNA." In: *Biochemical and Biophysical Research Communications* 421.4 (2012), pp. 671–677. DOI: 10.1016/j.bbrc.2012.04.059.

[3] Kerstin Andersson, Per Aake Malmqvist, and Björn O. Roos. "Second-order perturbation theory with a complete active space self-consistent field reference function." In: *The Journal of Chemical Physics* 96.2 (1992), pp. 1218–1226. DOI: 10.1063/1.462209.

[4] Kerstin. Andersson, Per Aake. Malmqvist, Bjo
rn O. Roos, Andrzej J. Sadlej, and Krzysztof. Wolinski. "Second-order perturbation theory with a CASSCF reference function." In: *The Journal of Physical Chemistry* 94.14 (1990), pp. 5483–5488. DOI: 10.1021/j100377a012.

[5] George B. Arfken. *Mathematical Methods for Physicists*. Elsevier Science & Techn., 2013. 1008 pp. ISBN: 9781483277820. URL: https://www.ebook.de/de/product/23273197/george_b_arfken_mathematical_methods_for_physicists.html.

[6] Anna Arola and Ramon Vilar. "Stabilisation of G-Quadruplex DNA by Small Molecules." In: *Current Topics in Medicinal Chemistry* 8.15 (2008), pp. 1405–1415. DOI: 10.2174/156802608786141106.

[7] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. García, and J.M. Soler. "Linear-Scaling ab-initio Calculations for Large and Complex Systems." In: *physica status solidi (b)* 215.1 (1999), pp. 809–817. DOI: 10.1002/(sici)1521-3951(199909)215:1<809::aid-pssb809>3.0.co;2-0.

[8] Alexander A. Auer and Marcel Nooijen. "Dynamically screened local correlation method using enveloping localized orbitals." In: *The Journal of Chemical Physics* 125.2 (2006), p. 024104. DOI: 10.1063/1.2209685.

[9] Philippe Y. Ayala and Gustavo E. Scuseria. "Atom pair partitioning of the correlation energy." In: *Chemical Physics Letters* 322.3-4 (2000), pp. 213–218. DOI: 10.1016/s0009-2614(00)00417-6.

[10] G. B. Bachelet, D. R. Hamann, and M. Schlüter. "Pseudopotentials that work: From H to Pu." In: *Physical Review B* 26.8 (1982), pp. 4199–4228. DOI: 10.1103/physrevb.26.4199.

[11] Giovanni B. Bachelet and M. Schlüter. "Relativistic norm-conserving pseudopotentials." In: *Physical Review B* 25.4 (1982), pp. 2103–2108. DOI: 10.1103/physrevb.25.2103.

[12] George B. Bacskay. "A quadratically convergent Hartree-Fock (QC-SCF) method. Application to closed shell systems." In: *Chemical Physics* 61.3 (1981), pp. 385–404. DOI: 10.1016/0301-0104(81)85156-7.

[13] R. F. W. Bader. "A Bond Path: A Universal Indicator of Bonded Interactions." In: *The Journal of Physical Chemistry A* 102.37 (1998), pp. 7314–7323. DOI: 10.1021/jp981794v.

[14] R. F. W. Bader and H. Essén. "The characterization of atomic interactions." In: *The Journal of Chemical Physics* 80.5 (1984), pp. 1943–1960. DOI: 10.1063/1.446956.

[15] Richard Bader. *Atoms in molecules : a quantum theory*. Oxford: Clarendon Press, 1990. ISBN: 0198551681.

[16] Bruce C. Baguley. "A brief history of cancer chemiotherapy." In: *Anticancer Drug Development*. Elsevier, 2002, pp. 1–11. DOI: 10.1016/b978-012072651-6/50002-4.

[17] Jon Baker. "An algorithm for the location of transition states." In: *Journal of Computational Chemistry* 7.4 (1986), pp. 385–395. DOI: 10.1002/jcc.540070402.

[18] Shankar Balasubramanian, Laurence H. Hurley, and Stephen Neidle. "Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?" In: *Nature Reviews Drug Discovery* 10.4 (2011), pp. 261–275. DOI: 10.1038/nrd3428.

[19] Pavel Banáš, Petr Jurečka, Nils G. Walter, Jiří Šponer, and Michal Otyepka. "Theoretical studies of RNA catalysis: Hybrid QM/MM methods and their comparison with MD and QM." In: *Methods* 49.2 (2009), pp. 202–216. DOI: 10.1016/j.ymeth.2009.04.007.

[20] Daniel Bandarra et al. "Mo(II) complexes: A new family of cytotoxic agents?" In: *Journal of Inorganic Biochemistry* 104.11 (2010), pp. 1171–1177. DOI: 10.1016/j.jinorgbio.2010.07.006.

[21] Amartya S. Banerjee, Phanish Suryanarayana, and John E. Pask. "Periodic Pulay method for robust and efficient convergence acceleration of self-consistent field iterations." In: *Chemical Physics Letters* 647 (2016), pp. 31–35. DOI: 10.1016/j.cplett.2016.01.033.

[22] Pier Giovanni Baraldi, Andrea Bovero, Francesca Fruttarolo, Delia Preti, Mojgan Aghazadeh Tabrizi, Maria Giovanna Pavani, and Romeo Romagnoli. "DNA minor groove binders as potential antitumor and antimicrobial agents." In: *Medicinal Research Reviews* 24.4 (2004), pp. 475–528. DOI: 10.1002/med.20000.

[23] Christopher M. Barbieri, Annankoil R. Srinivasan, Suzanne G. Rzuczek, Joseph E. Rice, Edmond J. LaVoie, and Daniel S. Pilch. "Defining the mode, energetics and specificity with which a macrocyclic hexaoxazole binds to human telomeric G-quadruplex DNA." In: *Nucleic Acids Research* 35.10 (2007), pp. 3272–3286. DOI: 10.1093/nar/gkm188.

[24] Rodney J. Bartlett. "Coupled-Cluster theory: An overview of recent developments." In: *Modern Electronic Structure Theory*. World Scientific Publishing Company, 1995, pp. 1047–1131. DOI: 10.1142/9789812832115_0005.

[25] Rodney J. Bartlett and Monika Musiał. "Coupled-cluster theory in quantum chemistry." In: *Reviews of Modern Physics* 79.1 (2007), pp. 291–352. DOI: 10.1103/revmodphys.79.291.

[26] Rodney J. Bartlett and George D. Purvis. "Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem." In: *International Journal of Quantum Chemistry* 14.5 (1978), pp. 561–581. DOI: 10.1002/qua.560140504.

[27] Jacqueline K. Barton, Douglas C. Rees, Clara L. Kielkopf, Kathryn E. Erkkila, and Brian P. Hudson. "Structure of a photoactive rhodium complex intercalated into DNA." In: *Nature Structural Biology* 7.2 (2000), pp. 117–121. DOI: 10.1038/72385.

[28] Hemanta Baruah, Colin G. Barry, and Ulrich Bierbach. "Platinum-Intercalator Conjugates: From DNA-Targeted Cisplatin Derivatives to Adenine Binding Complexes as Potential Modulators of Gene Regulation." In: *Current Topics in Medicinal Chemistry* 4.15 (2004), pp. 1537–1549. DOI: 10.2174/1568026043387313.

[29] J.-D. Beaudoin and J.-P. Perreault. "5-UTR G-quadruplex structures acting as translational repressors." In: *Nucleic Acids Research* 38.20 (2010), pp. 7022–7036. DOI: 10.1093/nar/gkq557.

[30] Axel D. Becke. "Density-functional thermochemistry. III. The role of exact exchange." In: *The Journal of Chemical Physics* 98.7 (1993), pp. 5648–5652. DOI: 10.1063/1.464913.

[31] S. Benito, A. Ferrer, S. Benabou, A. Aviñó, R. Eritja, and R. Gargallo. "Evaluation of the effect of polymorphism on G-quadruplex-ligand interaction by means of spectroscopic and chromatographic techniques." In: *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 196 (2018), pp. 185–195. DOI: 10.1016/j.saa.2018.02.006.

[32] Debmalya Bhattacharyya, Gayan Mirihana Arachchilage, and Soumitra Basu. "Metal Cations in G-Quadruplex Folding and Stability." In: *Frontiers in Chemistry* 4 (2016). DOI: 10.3389/fchem.2016.00038.

[33] F. Matthias Bickelhaupt and Evert Jan Baerends. "Kohn-Sham Density Functional Theory: Predicting and Understanding Chemistry." In: *Reviews in Computational Chemistry*. John Wiley & Sons, Inc., 2007, pp. 1–86. DOI: 10.1002/9780470125922.ch1.

[34] Giulia Biffi, David Tannahill, John McCafferty, and Shankar Balasubramanian. "Quantitative visualization of DNA G-quadruplex structures in human cells." In: *Nature Chemistry* 5.3 (2013), pp. 182–186. DOI: 10.1038/nchem.1548.

[35] Peter Blaha, Karlheinz Schwarz, Georg K.H. Madsen, Dieter Kvasnicka, Joachim Luitz, et al. "Wien2K: An augmented plane wave+ local orbitals program for calculating crystal properties." In: 60 (2001).

[36] Riccardo Bonsignore, Fabrizia Russo, Alessio Terenzi, Angelo Spinello, Antonino Lauria, Giuseppe Gennaro, Anna Maria Almerico, Bernhard K. Keppler, and Giampaolo Barone. "The interaction of Schiff Base complexes of nickel(II) and zinc(II) with duplex and G-quadruplex DNA." In: *Journal of Inorganic Biochemistry* 178 (2018), pp. 106–114. DOI: 10.1016/j.jinorgbio.2017.10.010.

[37] Riccardo Bonsignore, Alessio Terenzi, Angelo Spinello, Annamaria Martorana, Antonino Lauria, Anna Maria Almerico, Bernhard K. Keppler, and Giampaolo Barone. "G-quadruplex vs. duplex-DNA binding of nickel(II) and zinc(II) Schiff base complexes." In: *Journal of Inorganic Biochemistry* 161 (2016), pp. 115–121. DOI: 10.1016/j.jinorgbio.2016.05.010.

[38] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln." In: *Annalen der Physik* 389.20 (1927), pp. 457–484. DOI: 10.1002/andp.19273892002.

[39] James W. Boughton and Peter Pulay. "Comparison of the boys and Pipek-Mezey localizations in the local correlation approach and automatic virtual basis selection." In: *Journal of Computational Chemistry* 14.6 (1993), pp. 736–740. DOI: 10.1002/jcc.540140615.

[40] D R Bowler and T Miyazaki. "O(N) methods in electronic structure calculations." In: *Reports on Progress in Physics* 75.3 (2012), p. 036503. DOI: 10.1088/0034-4885/75/3/036503.

[41] Pathik S. Brahmkshatriya, Petr Dobes, Jindrich Fanfrlik, Jan Rezac, Kamil Paruch, Agnieszka Bronowska, Martin Lepsík, and Pavel Hobza. "Quantum Mechanical Scoring: Structural and Energetic Insights into Cyclin-Dependent Kinase 2 Inhibition by Pyrazolo[1,5-a]pyrimidines." In: *Current Computer Aided-Drug Design* 9.1 (2013), pp. 118–129. DOI: 10.2174/1573409911309010011.

[42] Michael S. Braun and Matthew T. Seymour. "Balancing the efficacy and toxicity of chemotherapy in colorectal cancer." In: *Therapeutic Advances in Medical Oncology* 3.1 (2010), pp. 43–52. DOI: 10.1177/1758834010388342.

[43] Craig R. Brodie, J. Grant Collins, and Janice R. Aldrich-Wright. "DNA binding and biological activity of some platinum(ii) intercalating compounds containing methyl-substituted 1,10-phenanthrolines." In: *Dalton Transactions* 8 (2004), p. 1145. DOI: 10.1039/b316511f.

[44] Sarah Burge, Gary N. Parkinson, Pascale Hazel, Alan K. Todd, and Stephen Neidle. "Quadruplex DNA: sequence, topology and structure." In: *Nucleic Acids Research* 34.19 (2006), pp. 5402–5415. DOI: 10.1093/nar/gkl655.

[45] Angelika M. Burger, Fangping Dai, Christoph M. Schultes, Anthony P. Reszka, Michael J. Moore, John A. Double, and Stephen Neidle. "The G-Quadruplex-Interactive Molecule BRACO-19 Inhibits Tumor Growth, Consistent with Telomere Targeting and Interference with Telomerase Function." In: *Cancer Research* 65.4 (2005), pp. 1489–1496. DOI: 10.1158/0008-5472.can-04-2910.

[46] CONQUEST:http://www.order-n.org/.

[47] Ross Cagan and Pablo Meyer. "Rethinking cancer: current challenges and opportunities in cancer research." In: *Disease Models & Mechanisms* 10.4 (2017), pp. 349–352. DOI: 10.1242/dmm.030007.

[48] Qian Cao, Yi Li, Eva Freisinger, Peter Z. Qin, Roland K. O. Sigel, and Zong-Wan Mao. "G-quadruplex DNA targeted metal complexes acting as potential anticancer drugs." In: *Inorganic Chemistry Frontiers* 4.1 (2017), pp. 10–32. DOI: 10.1039/c6qi00300a.

[49] Anders S. Christensen, Tomáš Kubař, Qiang Cui, and Marcus Elstner. "Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications." In: *Chemical Reviews* 116.9 (2016), pp. 5301–5337. DOI: 10.1021/acs.chemrev.5b00584.

[50] Ove Christiansen, Henrik Koch, and Poul Jørgensen. "The second-order approximate coupled cluster singles and doubles model CC2." In: *Chemical Physics Letters* 243.5-6 (1995), pp. 409–418. DOI: 10.1016/0009-2614(95)00841-q.

[51] Wei Chua, Patricia S. Kho, Melissa M. Moore, Kellie A. Charles, and Stephen J. Clarke. "Clinical, laboratory and molecular factors predicting chemotherapy efficacy and toxicity in colorectal cancer." In: *Critical Reviews in Oncology/Hematology* 79.3 (2011), pp. 224–250. DOI: 10.1016/j.critrevonc.2010.07.012.

[52] Gwo-Yu Chuang, Dima Kozakov, Ryan Brenke, Stephen R. Comeau, and Sandor Vajda. "DARS (Decoys As the Reference State) Potentials for Protein-Protein Docking." In: *Biophysical Journal* 95.9 (2008), pp. 4217–4227. DOI: 10.1529/biophysj.108.135814.

[53] Wan Jun Chung, Brahim Heddi, Emmanuelle Schmitt, Kah Wai Lim, Yves Mechulam, and Anh Tuân Phan. "Structure of a left-handed DNA G-quadruplex." In: *Proceedings of the National Academy of Sciences* 112.9 (2015), pp. 2729–2733. DOI: 10.1073/pnas.1418718112.

[54] J Cizek and J Paldus. "Coupled Cluster Approach." In: *Physica Scripta* 21.3-4 (1980), pp. 251–254. DOI: 10.1088/0031-8949/21/3-4/006.

[55] W.T. Cochran, J.W. Cooley, D.L. Favin, H.D. Helms, R.A. Kaenel, W.W. Lang, G.C. Maling, D.E. Nelson, C.M. Rader, and P.D. Welch. "What is the fast Fourier transform?" In: *Proceedings of the IEEE* 55.10 (1967), pp. 1664–1674. DOI: 10.1109/proc.1967.5957.

[56] Gavin W. Collie, Gary N. Parkinson, Stephen Neidle, Frédéric Rosu, Edwin De Pauw, and Valérie Gabelica. "Electrospray Mass Spectrometry of Telomeric RNA (TERRA) Reveals the Formation of Stable Multimeric G-Quadruplex Structures." In: *Journal of the American Chemical Society* 132.27 (2010), pp. 9328–9334. DOI: 10.1021/ja100345z.

[57] Julia Contreras-García, Erin R. Johnson, Shahar Keinan, Robin Chaudret, Jean-Philip Piquemal, David N. Beratan, and Weitao Yang. "NCIPLOT: A Program for Plotting Noncovalent Interac-

tion Regions." In: *Journal of Chemical Theory and Computation* 7.3 (2011), pp. 625–632. DOI: 10.1021/ct100641a.

[58] J.J. Dannenberg. "Hydrogen bonds: a comparison of semiempirical and ab initio treatments." In: *Journal of Molecular Structure: THEOCHEM* 401.3 (1997), pp. 279–286. DOI: 10.1016/s0166-1280(97)00029-8.

[59] William A. Denny. "DNA Minor Groove Alkylating Agents." In: *Current Medicinal Chemistry* 8.5 (2001), pp. 533–544. DOI: 10.2174/0929867003373283.

[60] Haluk Dinçalp, Şevki Kızılok, Özgül Haklı Birel, and Sıddık İçli. "Synthesis and G-quadruplex binding study of a novel full visible absorbing perylene diimide dye." In: *Journal of Photochemistry and Photobiology A: Chemistry* 235 (2012), pp. 40–48. DOI: 10.1016/j.jphotochem.2012.03.004.

[61] Paul Adrien Maurice Dirac. "On the theory of quantum mechanics." In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 112.762 (1926), pp. 661–677. DOI: 10.1098/rspa.1926.0133.

[62] Isabelle M. Dixon, Frédéric Lopez, Agueda M. Tejera, Jean-Pierre Estève, Maria A. Blasco, Geneviève Pratviel, and Bernard Meunier. "A G-Quadruplex Ligand with 10000-Fold Selectivity over Duplex DNA." In: *Journal of the American Chemical Society* 129.6 (2007), pp. 1502–1503. DOI: 10.1021/ja065591t.

[63] ERGOSCF:http://www.ergoscf.org/.

[64] A.W. Ehlers, M. Böhme, S. Dapprich, A. Gobbi, A. Höllwarth, V. Jonas, K.F. Köhler, R. Stegmann, A. Veldkamp, and G. Frenking. "A set of f-polarization functions for pseudo-potential basis sets of the transition metals Sc-Cu, Y-Ag and La-Au." In: *Chemical Physics Letters* 208.1-2 (1993), pp. 111–114. DOI: 10.1016/0009-2614(93)80086-5.

[65] Sawssen Elleuchi, Iker Ortiz de Luzuriaga, Ángel Sanchez-Gonzalez, Xabier Lopez, Khaled Jarraya, Maria José Calhorda, and Adrià Gil. "Computational Studies on the Binding Preferences of Molybdenum(II) Phenanthroline Complexes with Duplex DNA. The Important Role of the Ancillary Ligands." In:

Inorganic Chemistry 59.17 (2020), pp. 12711–12721. DOI: 10.1021/acs.inorgchem.0c01793.

[66] Eduardo V. Ludeña Eugene S. Kryachko. *Energy Density Functional Theory of Many-Electron Systems*. Springer Netherlands, 2012. 850 pp. ISBN: 9789400919709. URL: https://www.ebook.de/de/product/25181670/eugene_s_kryachko_eduardo_v_ludena_energy_density_functional_theory_of_many_electron_systems.html.

[67] FREEON:https://launchpad.net/freeon.

[68] Gianluca Farine, Claudio Migliore, Alessio Terenzi, Fabrizio Lo Celso, Antonio Santoro, Giuseppe Bruno, Riccardo Bonsignore, and Giampaolo Barone. "On the G-Quadruplex Binding of a New Class of Nickel(II), Copper(II), and Zinc(II) Salphen-Like Complexes." In: *European Journal of Inorganic Chemistry* 2021.14 (2021), pp. 1332–1336. DOI: 10.1002/ejic.202100067.

[69] John C. Faver, Mark L. Benson, Xiao He, Benjamin P. Roberts, Bing Wang, Michael S. Marshall, Matthew R. Kennedy, C. David Sherrill, and Kenneth M. Merz. "Formal Estimation of Errors in Computed Absolute Interaction Energies of Protein-Ligand Complexes." In: *Journal of Chemical Theory and Computation* 7.3 (2011), pp. 790–797. DOI: 10.1021/ct100563b.

[70] Yun Feng, Dazhang Yang, Hongbo Chen, Wenli Cheng, Lixia Wang, Hongxia Sun, and Yalin Tang. "Stabilization of G-quadruplex DNA and inhibition of Bcl-2 expression by a pyridostatin analog." In: *Bioorganic & Medicinal Chemistry Letters* 26.7 (2016), pp. 1660–1663. DOI: 10.1016/j.bmcl.2016.02.065.

[71] Terace M. Fletcher, Daekyu Sun, Miguel Salazar, and Laurence H. Hurley. "Effect of DNA Secondary Structure on Human Telomerase Activity." In: *Biochemistry* 37.16 (1998), pp. 5536–5541. DOI: 10.1021/bi972681p.

[72] V. Fock. "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems." In: *Zeitschrift für Physik* 61.1-2 (1930), pp. 126–148. DOI: 10.1007/bf01340294.

[73] M. J. Frisch et al. *Gaussian16 Revision C.01*. 2016.

[74] Rodrigo Galindo-Murillo, James C. Robertson, Marie Zgarbová, Jiří Šponer, Michal Otyepka, Petr Jurečka, and Thomas E. Cheatham. "Assessing the Current State of Amber Force Field Modifications for DNA." In: *Journal of Chemical Theory and Computation* 12.8 (2016), pp. 4114–4127. DOI: 10.1021/acs.jctc.6b00186.

[75] Aurellia Galliot, Adrià Gil, and Maria José Calhorda. "Effects of oxygenation on the intercalation of 1,10-phenanthroline-5,6/4,7-dione between DNA base pairs: a computational study." In: *Physical Chemistry Chemical Physics* 19.25 (2017), pp. 16638–16649. DOI: 10.1039/c7cp00532f.

[76] Alberto García, Matthieu J. Verstraete, Yann Pouillon, and Javier Junquera. "The psml format and library for norm-conserving pseudopotential data curation and interoperability." In: *Computer Physics Communications* 227 (2018), pp. 51–71. DOI: 10.1016/j.cpc.2018.02.011.

[77] A. Gil, V. Branchadell, and M. J. Calhorda. "A theoretical study of methylation and CH/π interactions in DNA intercalation: methylated 1,10-phenanthroline in adenine-thymine base pairs." In: *RSC Advances* 6.89 (2016), pp. 85891–85902. DOI: 10.1039/c6ra15495f.

[78] Adrià Gil, Manuel Melle-Franco, Vicenç Branchadell, and Maria José Calhorda. "How the Intercalation of Phenanthroline Affects the Structure, Energetics, and Bond Properties of DNA Base Pairs: Theoretical Study Applied to Adenine-Thymine and Guanine-Cytosine Tetramers." In: *Journal of Chemical Theory and Computation* 11.6 (2015), pp. 2714–2728. DOI: 10.1021/ct5006104.

[79] Adrià Gil, Ángel Sánchez-González, and Vicenç Branchadell. "Unraveling the Modulation of the Activity in Drugs Based on Methylated Phenanthroline When Intercalating between DNA Base Pairs." In: *Journal of Chemical Information and Modeling* 59.9 (2019), pp. 3989–3995. DOI: 10.1021/acs.jcim.9b00500.

[80] Konstantinos Gkionis, Holger Kruse, James A. Platts, Arnošt Mládek, Jaroslav Koča, and Jiří Šponer. "Ion Binding to Quadruplex DNA Stems. Comparison of MM and QM Descriptions Reveals Sizable Polarization Effects Not Included in Contempo-

rary Simulations." In: *Journal of Chemical Theory and Computation* 10.3 (2014), pp. 1326–1340. DOI: 10.1021/ct4009969.

[81] Stefan Goedecker. "Linear scaling electronic structure methods." In: *Reviews of Modern Physics* 71.4 (1999), pp. 1085–1123. DOI: 10.1103/revmodphys.71.1085.

[82] Stefan Grimme. "Accurate description of van der Waals complexes by density functional theory including empirical corrections." In: *Journal of Computational Chemistry* 25.12 (2004), pp. 1463–1473. DOI: 10.1002/jcc.20078.

[83] Stefan Grimme. "Semiempirical GGA-type density functional constructed with a long-range dispersion correction." In: *Journal of Computational Chemistry* 27.15 (2006), pp. 1787–1799. DOI: 10.1002/jcc.20495.

[84] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu." In: *The Journal of Chemical Physics* 132.15 (2010), p. 154104. DOI: 10.1063/1.3382344.

[85] Jörg Grunenberg, Giampaolo Barone, and Angelo Spinello. "The Right Answer for the Right Electrostatics: Force Field Methods Are Able to Describe Relative Energies of DNA Guanine Quadruplexes." In: *Journal of Chemical Theory and Computation* 10.8 (2014), pp. 2901–2905. DOI: 10.1021/ct500329f.

[86] Federica Guarra, Tiziano Marzo, Marta Ferraroni, Francesco Papi, Carla Bazzicalupi, Paola Gratteri, Gennaro Pescitelli, Luigi Messori, Tarita Biver, and Chiara Gabbiani. "Interaction of a gold(i) dicarbene anticancer drug with human telomeric DNA G-quadruplex: solution and computationally aided X-ray diffraction analysis." In: *Dalton Transactions* 47.45 (2018), pp. 16132–16138. DOI: 10.1039/c8dt03607a.

[87] C. Fonseca Guerra, J. G. Snijders, G. te Velde, and E. J. Baerends. "Towards an order-N DFT method." In: *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* 99.6 (1998), pp. 391–403. DOI: 10.1007/s002140050353.

[88] Célia Fonseca Guerra, Tushar van der Wijst, Jordi Poater, Marcel Swart, and F. Matthias Bickelhaupt. "Adenine versus guanine quartets in aqueous solution: dispersion-corrected DFT study on the differences in π-stacking and hydrogen-bonding behavior." In: *Theoretical Chemistry Accounts* 125.3-6 (2009), pp. 245–252. DOI: 10.1007/s00214-009-0634-9.

[89] Célia Fonseca Guerra, Hester Zijlstra, Gábor Paragi, and F. Matthias Bickelhaupt. "Telomere Structure and Stability: Covalency in Hydrogen Bonds, Not Resonance Assistance, Causes Cooperativity in Guanine Quartets." In: *Chemistry - A European Journal* 17.45 (2011), pp. 12612–12622. DOI: 10.1002/chem.201102234.

[90] Shozeb M. Haider, Gary N. Parkinson, and Stephen Neidle. "Structure of a G-quadruplex-Ligand Complex." In: *Journal of Molecular Biology* 326.1 (2003), pp. 117–125. DOI: 10.1016/s0022-2836(02)01354-2.

[91] Shozeb Haider, Gary N. Parkinson, and Stephen Neidle. "Crystal Structure of the Potassium Form of an Oxytricha nova G-quadruplex." In: *Journal of Molecular Biology* 320.2 (2002), pp. 189–200. DOI: 10.1016/s0022-2836(02)00428-x.

[92] Shozeb Haider, Gary N. Parkinson, and Stephen Neidle. "Molecular Dynamics and Principal Components Analysis of Human Telomeric Quadruplex Multimers." In: *Biophysical Journal* 95.1 (2008), pp. 296–311. DOI: 10.1529/biophysj.107.120501.

[93] D. R. Hamann, M. Schlüter, and C. Chiang. "Norm-Conserving Pseudopotentials." In: *Physical Review Letters* 43.20 (1979), pp. 1494–1497. DOI: 10.1103/physrevlett.43.1494.

[94] Paris L. Hamilton and Dev P. Arya. "Natural product DNA major groove binders." In: *Nat. Prod. Rep.* 29.2 (2012), pp. 134–143. DOI: 10.1039/c1np00054c.

[95] Claudia Hampel and Hans-Joachim Werner. "Local treatment of electron correlation in coupled cluster theory." In: *The Journal of Chemical Physics* 104.16 (1996), pp. 6286–6297. DOI: 10.1063/1.471289.

[96] K.R. Harrap. "Preclinical studies identifying carboplatin as a viable cisplatin alternative." In: *Cancer Treatment Reviews* 12 (1985), pp. 21–33. DOI: 10.1016/0305-7372(85)90015-5.

[97] D. R. Hartree. "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods." In: *Mathematical Proceedings of the Cambridge Philosophical Society* 24.1 (1928), pp. 89–110. DOI: 10.1017/s0305004100011919.

[98] Marek Havrila, Petr Stadlbauer, Barira Islam, Michal Otyepka, and Jiří Šponer. "Effect of Monovalent Ion Parameters on Molecular Dynamics Simulations of G-Quadruplexes." In: *Journal of Chemical Theory and Computation* 13.8 (2017), pp. 3911–3926. DOI: 10.1021/acs.jctc.7b00257.

[99] P. Jeffrey Hay and Willard R. Wadt. "Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals." In: *The Journal of Chemical Physics* 82.1 (1985), pp. 299–310. DOI: 10.1063/1.448975.

[100] Lei He, Xiang Chen, Zhenyu Meng, Jintao Wang, Keyin Tian, Tianhu Li, and Fangwei Shao. "Octahedral ruthenium complexes selectively stabilize G-quadruplexes." In: *Chemical Communications* 52.52 (2016), pp. 8095–8098. DOI: 10.1039/c6cc03117j.

[101] Werner Heisenberg. "Quantum-theoretical re-interpretation of kinematic and mechanical relations." In: *Z. Phys* (1925).

[102] Junming Ho, Michael B. Newcomer, Christina M. Ragain, Jose A. Gascon, Enrique R. Batista, J. Patrick Loria, and Victor S. Batista. "MoD-QM/MM Structural Refinement Method: Characterization of Hydrogen Bonding in the Oxytricha nova G-Quadruplex." In: *Journal of Chemical Theory and Computation* 10.11 (2014), pp. 5125–5135. DOI: 10.1021/ct500571k.

[103] Yee-Ping Ho, Steve C.F. Au-Yeung, and Kenneth K.W. To. "Platinum-based anticancer agents: Innovative design strategies and biological perspectives." In: *Medicinal Research Reviews* 23.5 (2003), pp. 633–655. DOI: 10.1002/med.10038.

[104] D. Hobbs, G. Kresse, and J. Hafner. "Fully unconstrained noncollinear magnetism within the projector augmented-wave method." In: *Physical Review B* 62.17 (2000), pp. 11556–11570. DOI: 10.1103/physrevb.62.11556.

[105] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas." In: *Physical Review* 136.3B (1964), B864–B871. DOI: 10.1103/physrev.136.b864.

[106] Moritz von Hopffgarten and Gernot Frenking. "Energy decomposition analysis." In: *WIREs Computational Molecular Science* 2.1 (2011), pp. 43–62. DOI: 10.1002/wcms.71.

[107] Jiří Hostaš, Jan Řezáč, and Pavel Hobza. "On the performance of the semiempirical quantum mechanical PM6 and PM7 methods for noncovalent interactions." In: *Chemical Physics Letters* 568-569 (2013), pp. 161–166. DOI: 10.1016/j.cplett.2013.02.069.

[108] Candide Hounsou, Lionel Guittat, David Monchaud, Muriel Jourdan, Nicolas Saettel, Jean-Louis Mergny, and Marie-Paule Teulade-Fichou. "G-Quadruplex Recognition by Quinacridines: a SAR, NMR, and Biological Study." In: *ChemMedChem* 2.5 (2007), pp. 655–666. DOI: 10.1002/cmdc.200600286.

[109] Nicholas V. Hud, Flint W. Smith, Frank A. L. Anet, and Juli Feigon. "The Selectivity for K^+ versus Na^+ in DNA Quadruplexes Is Dominated by Relative Free Energies of Hydration: A Thermodynamic Analysis by ^{1}H NMR." In: *Biochemistry* 35.48 (1996), pp. 15383–15390. DOI: 10.1021/bi9620565.

[110] Linda Hung, Chen Huang, Ilgyou Shin, Gregory S. Ho, Vincent L. Lignères, and Emily A. Carter. "Introducing PROFESS 2.0: A parallelized, fully linear scaling program for orbital-free density functional theory calculations." In: *Computer Physics Communications* 181.12 (2010), pp. 2208–2209. DOI: 10.1016/j.cpc.2010.09.001.

[111] Robert Hänsel-Hertsch et al. "G-quadruplex structures mark human regulatory chromatin." In: *Nature Genetics* 48.10 (2016), pp. 1267–1272. DOI: 10.1038/ng.3662.

[112] Barira Islam, Miriam Sgobba, Charlie Laughton, Modesto Orozco, Jiri Sponer, Stephen Neidle, and Shozeb Haider. "Conformational dynamics of the human propeller telomeric DNA quadruplex on a microsecond time scale." In: *Nucleic Acids Research* 41.4 (2013), pp. 2723–2735. DOI: 10.1093/nar/gks1331.

[113] Barira Islam, Petr Stadlbauer, Miroslav Krepl, Marek Havrila, Shozeb Haider, and Jiri Sponer. "Structural Dynamics of Lateral and Diagonal Loops of Human Telomeric G-Quadruplexes in Extended MD Simulations." In: *Journal of Chemical Theory and Computation* 14.10 (2018), pp. 5011–5026. DOI: 10.1021/acs.jctc.8b00543.

[114] Barira Islam, Petr Stadlbauer, Michaela Vorlíčková, Jean-Louis Mergny, Michal Otyepka, and Jiří Šponer. "Stability of Two-Quartet G-Quadruplexes and Their Dimers in Atomistic Simulations." In: *Journal of Chemical Theory and Computation* 16.6 (2020), pp. 3447–3463. DOI: 10.1021/acs.jctc.9b01068.

[115] Ivan Ivani et al. "Parmbsc1: a refined force field for DNA simulations." In: *Nature Methods* 13.1 (2015), pp. 55–58. DOI: 10.1038/nmeth.3658.

[116] Benny G. Johnson, Peter M. W. Gill, and John A. Pople. "The performance of a family of density functional methods." In: *The Journal of Chemical Physics* 98.7 (1993), pp. 5612–5626. DOI: 10.1063/1.464906.

[117] D. D. Johnson. "Modified Broyden's method for accelerating convergence in self-consistent calculations." In: *Physical Review B* 38.18 (1988), pp. 12807–12813. DOI: 10.1103/physrevb.38.12807.

[118] Erin R. Johnson, Shahar Keinan, Paula Mori-Sánchez, Julia Contreras-García, Aron J. Cohen, and Weitao Yang. "Revealing Noncovalent Interactions." In: *Journal of the American Chemical Society* 132.18 (2010), pp. 6498–6506. DOI: 10.1021/ja100936w.

[119] Javier Junquera, Óscar Paz, Daniel Sánchez-Portal, and Emilio Artacho. "Numerical atomic orbitals for linear-scaling calculations." In: *Physical Review B* 64.23 (2001), p. 235111. DOI: 10.1103/physrevb.64.235111.

[120] Petr Jurečka, Jiří Šponer, Jiří Černý, and Pavel Hobza. "Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs." In: *Phys. Chem. Chem. Phys.* 8.17 (2006), pp. 1985–1993. DOI: 10.1039/b600027d.

[121] Nurul H. Abd Karim, Oscar Mendoza, Arun Shivalingam, Alexander J. Thompson, Sushobhan Ghosh, Marina K. Kuimova, and Ramon Vilar. "Salphen metal complexes as tunable G-quadruplex binders and optical probes." In: *RSC Adv.* 4.7 (2014), pp. 3355–3363. DOI: 10.1039/c3ra44793f.

[122] Todd A. Keith. *AIMAll (Version 19.10.12), TK Gristmill Software, Overland Park KS, USA*. 2019.

[123] Andrew Kellett, Zara Molphy, Creina Slator, Vickie McKee, and Nicholas P. Farrell. "Molecular methods for assessment of non-covalent metallodrug-DNA interactions." In: *Chemical Society Reviews* 48.4 (2019), pp. 971–988. DOI: 10.1039/c8cs00157j.

[124] Roxanne Kieltyka, Johans Fakhoury, Nicolas Moitessier, and Hanadi F. Sleiman. "Platinum Phenanthroimidazole Complexes as G-Quadruplex DNA Selective Binders." In: *Chemistry - A European Journal* 14.4 (2008), pp. 1145–1154. DOI: 10.1002/chem.200700783.

[125] Mu-Yong Kim, Hariprasad Vankayalapati, Kazuo Shin-ya, Konstanty Wierzba, and Laurence H. Hurley. "Telomestatin, a Potent Telomerase Inhibitor That Interacts Quite Specifically with the Human Telomeric Intramolecular G-Quadruplex." In: *Journal of the American Chemical Society* 124.10 (2002), pp. 2098–2099. DOI: 10.1021/ja017308q.

[126] Kazuo Kitaura and Keiji Morokuma. "A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation." In: *International Journal of Quantum Chemistry* 10.2 (1976), pp. 325–340. DOI: 10.1002/qua.560100211.

[127] A. Klamt and G. Schüürmann. "COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient." In: *J. Chem. Soc., Perkin Trans. 2* 5 (1993), pp. 799–805. DOI: 10.1039/p29930000799.

[128] Leonard Kleinman. "Relativistic norm-conserving pseudopotential." In: *Physical Review B* 21.6 (1980), pp. 2630–2631. DOI: 10.1103/physrevb.21.2630.

[129] Leonard Kleinman and D. M. Bylander. "Efficacious Form for Model Pseudopotentials." In: *Physical Review Letters* 48.20 (1982), pp. 1425–1428. DOI: 10.1103/physrevlett.48.1425.

[130] Peter J. Knowles and Hans-Joachim Werner. "Internally contracted multiconfiguration-reference configuration interaction calculations for excited states." In: *Theoretica Chimica Acta* 84.1-2 (1992), pp. 95–103. DOI: 10.1007/bf01117405.

[131] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects." In: *Physical Review* 140.4A (1965), A1133–A1138. DOI: 10.1103/physrev.140.a1133.

[132] Sofia Kolesnikova, Martin Hubálek, Lucie Bednárová, Josef Cvačka, and Edward A. Curtis. "Multimerization rules for G-quadruplexes." In: *Nucleic Acids Research* 45.15 (2017), pp. 8684–8696. DOI: 10.1093/nar/gkx637.

[133] Martin Korth. "Third-Generation Hydrogen-Bonding Corrections for Semiempirical QM Methods and Force Fields." In: *Journal of Chemical Theory and Computation* 6.12 (2010), pp. 3808–3816. DOI: 10.1021/ct100408b.

[134] Martin Korth. "Empirical Hydrogen-Bond Potential Functions-An Old Hat Reconditioned." In: *ChemPhysChem* 12.17 (2011), pp. 3131–3142. DOI: 10.1002/cphc.201100540.

[135] Martin Korth, Michal Pitoňák, Jan Řezáč, and Pavel Hobza. "A Transferable H-Bonding Correction for Semiempirical Quantum-Chemical Methods." In: *Journal of Chemical Theory and Computation* 6.1 (2009), pp. 344–352. DOI: 10.1021/ct900541n.

[136] Simone Kossmann and Frank Neese. "Comparison of two efficient approximate Hartee-Fock approaches." In: *Chemical Physics Letters* 481.4-6 (2009), pp. 240–243. DOI: 10.1016/j.cplett.2009.09.073.

[137] Irena Kostova. "Ruthenium Complexes as Anticancer Agents." In: *Current Medicinal Chemistry* 13.9 (2006), pp. 1085–1107. DOI: 10.2174/092986706776360941.

[138] G. Kresse and J. Furthmüller. "Efficient iterative schemes forab initiototal-energy calculations using a plane-wave basis set." In: *Physical Review B* 54.16 (1996), pp. 11169–11186. DOI: 10.1103/physrevb.54.11169.

[139] G. Kresse and D. Joubert. "From ultrasoft pseudopotentials to the projector augmented-wave method." In: *Physical Review B* 59.3 (1999), pp. 1758–1775. DOI: 10.1103/physrevb.59.1758.

[140] Jan K. Labanowski and Jan W Andzelm, eds. *Density Functional Methods in Chemistry*. Springer Nature, 1991. 443 pp. ISBN: 0387975128. URL: https://www.ebook.de/de/product/14592335/density_functional_methods_in_chemistry.html.

[141] Eric Largy, Jean-Louis Mergny, and Valérie Gabelica. "Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability." In: *The Alkali Metal Ions: Their Role for Life*. Springer International Publishing, 2016, pp. 203–258. DOI: 10.1007/978-3-319-21756-7_7.

[142] Anna Y. Lebedeva, Olga A. Fedorova, Vladimir B. Tsvetkov, Valerij Y. Grinberg, Natalia V. Grinberg, Tatiana V. Burova, Alexander S. Dubovik, Kirill K. Babievsky, and Yury V. Fedorov. "Novel 18-crown-6-ether containing mono- and bisstyryl dyes derived from pyridine moiety as fluorescent dyes for noncovalent interaction with DNA." In: *Dyes and Pigments* 157 (2018), pp. 80–92. DOI: 10.1016/j.dyepig.2018.04.038.

[143] Chengteh Lee, Weitao Yang, and Robert G. Parr. "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density." In: *Physical Review B* 37.2 (1988), pp. 785–789. DOI: 10.1103/physrevb.37.785.

[144] Kyuho Lee, Éamonn D. Murray, Lingzhu Kong, Bengt I. Lundqvist, and David C. Langreth. "Higher-accuracy van der Waals density functional." In: *Physical Review B* 82.8 (2010), p. 081101. DOI: 10.1103/physrevb.82.081101.

[145] Wook Lee and Spiridoula Matsika. "Conformational and electronic effects on the formation of anti cyclobutane pyrimidine dimers in G-quadruplex structures." In: *Physical Chemistry Chemical Physics* 19.4 (2017), pp. 3325–3336. DOI: 10.1039/c6cp05604k.

[146] E. van Lenthe, E. J. Baerends, and J. G. Snijders. "Relativistic regular two-component Hamiltonians." In: *The Journal of Chemical Physics* 99.6 (1993), pp. 4597–4610. DOI: 10.1063/1.466059.

[147] E. van Lenthe, E. J. Baerends, and J. G. Snijders. "Relativistic total energy using regular approximations." In: *The Journal of Chemical Physics* 101.11 (1994), pp. 9783–9792. DOI: 10.1063/1.467943.

[148] E. van Lenthe, R. van Leeuwen, E. J. Baerends, and J. G. Snijders. "Relativistic regular two-component Hamiltonians." In: *International Journal of Quantum Chemistry* 57.3 (1996), pp. 281–293. DOI: 10.1002/(sici)1097-461x(1996)57:3<281::aid-qua2>3.0.co;2-u.

[149] E. van Lenthe, J. G. Snijders, and E. J. Baerends. "The zero-order regular approximation for relativistic effects: The effect of spin-orbit coupling in closed shell molecules." In: *The Journal of Chemical Physics* 105.15 (1996), pp. 6505–6516. DOI: 10.1063/1.472460.

[150] Erik van Lenthe, Andreas Ehlers, and Evert-Jan Baerends. "Geometry optimizations in the zero order regular approximation for relativistic effects." In: *The Journal of Chemical Physics* 110.18 (1999), pp. 8943–8953. DOI: 10.1063/1.478813.

[151] Ka-Ho Leung, Victor Pui-Yan Ma, Hong-Zhang He, Daniel Shiu-Hin Chan, Hui Yang, Chung-Hang Leung, and Dik-Lung Ma. "A highly selective G-quadruplex-based luminescent switch-on probe for the detection of nanomolar strontium(ii) ions in sea water." In: *RSC Advances* 2.22 (2012), p. 8273. DOI: 10.1039/c2ra21119j.

[152] Guoliang Liao, Xiang Chen, Jingheng Wu, Chen Qian, Hanqiang Wang, Liangnian Ji, and Hui Chao. "Novel ruthenium(ii) polypyridyl complexes as G-quadruplex stabilisers and telomerase inhibitors." In: *Dalton Transactions* 43.21 (2014), p. 7811. DOI: 10.1039/c3dt53547a.

[153] Hai Lin and Donald G. Truhlar. "QM/MM: what have we learned, where are we, and where do we go from here?" In: *Theoretical Chemistry Accounts* 117.2 (2006). DOI: 10.1007/s00214-006-0143-z.

[154] Iker Ortiz de Luzuriaga, Xabier Lopez, and Adrià Gil. "Learning to Model G-Quadruplexes: Current Methods and Perspectives." In: *Annual Review of Biophysics* 50.1 (2021), pp. 209–243. DOI: 10.1146/annurev-biophys-060320-091827.

[155] Iker Ortiz de Luzuriaga, Ángel Sánchez-González, Wojciech Synoradzki, Xabier Lopez, and Adrià Gil. "Unravelling the binding affinity and selectivity of molybdenum(scpii/scp) phenanthroline complexes with DNA G-quadruplexes by using linear-scaling DFT studies. The important role of ancillary ligands." In: *Physical Chemistry Chemical Physics* 24.42 (2022), pp. 25918–25929. DOI: 10.1039/d2cp02241a.

[156] Georg K. H. Madsen, Peter Blaha, Karlheinz Schwarz, Elisabeth Sjöstedt, and Lars Nordström. "Efficient linearization of the augmented plane-wave method." In: *Physical Review B* 64.19 (2001), p. 195134. DOI: 10.1103/physrevb.64.195134.

[157] Rita P. Magalhães, Henriques S. Fernandes, and Sérgio F. Sousa. "Modelling Enzymatic Mechanisms with QM/MM Approaches: Current Status and Future Challenges." In: *Israel Journal of Chemistry* 60.7 (2020), pp. 655–666. DOI: 10.1002/ijch.202000014.

[158] Marta Matos, Carlos Romao, Claudia Cristina Lage Pereira, Sandra S. Rodrigues, Marcia Mora, Maria Joao Pires Silva, Paula M. Alves, and Celso Albuquerque Reis. *COMPOSITIONS COMPRISING ORGANOMETALLIC MOLYBDENUM COMPOUNDS FOR TREATING CANCER*. 2004.

[159] J.-L. Mergny. "Natural and pharmacological regulation of telomerase." In: *Nucleic Acids Research* 30.4 (2002), pp. 839–865. DOI: 10.1093/nar/30.4.839.

[160] Yoko Mikami-Terao, Masaharu Akiyama, Yuki Yuza, Takaaki Yanagisawa, Osamu Yamada, and Hisashi Yamada. "Antitumor activity of G-quadruplex-interactive agent TMPyP4 in K562 leukemic cells." In: *Cancer Letters* 261.2 (2008), pp. 226–234. DOI: 10.1016/j.canlet.2007.11.017.

[161] Gurpreet Singh Minhas, Daniel S. Pilch, John E. Kerrigan, Edmond J. LaVoie, and Joseph E. Rice. "Synthesis and G-quadruplex stabilizing properties of a series of oxazole-containing macrocycles." In: *Bioorganic & Medicinal Chemistry*

Letters 16.15 (2006), pp. 3891–3895. DOI: 10.1016/j.bmcl.2006.05.038.

[162] A. V. Mitin. "The dynamic "level shift" method for improving the convergence of the SCF procedure." In: *Journal of Computational Chemistry* 9.2 (1988), pp. 107–110. DOI: 10.1002/jcc.540090203.

[163] Kiana Gholamjani Moghaddam, Goran Giudetti, Wouter Sipma, and Shirin Faraji. "Theoretical insights into the effect of size and substitution patterns of azobenzene derivatives on the DNA G-quadruplex." In: *Physical Chemistry Chemical Physics* 22.46 (2020), pp. 26944–26954. DOI: 10.1039/d0cp04392c.

[164] Chr. Møller and M. S. Plesset. "Note on an Approximation Treatment for Many-Electron Systems." In: *Physical Review* 46.7 (1934), pp. 618–622. DOI: 10.1103/physrev.46.618.

[165] Keiji Morokuma. "Molecular Orbital Studies of Hydrogen Bonds. III C=O···H–O Hydrogen Bond in $H_2CO\cdots H_2O$ and $H_2CO\cdots 2H_2O$." In: *The Journal of Chemical Physics* ().

[166] Diana Mustard and David W. Ritchie. "Docking essential dynamics eigenstructures." In: *Proteins: Structure, Function, and Bioinformatics* 60.2 (2005), pp. 269–274. DOI: 10.1002/prot.20569.

[167] Frank Neese. "The ORCA program system." In: *WIREs Computational Molecular Science* 2.1 (2011), pp. 73–78. DOI: 10.1002/wcms.81.

[168] Frank Neese, Frank Wennmohs, and Andreas Hansen. "Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method." In: *The Journal of Chemical Physics* 130.11 (2009), p. 114108. DOI: 10.1063/1.3086717.

[169] Stephen Neidle. *Cancer Drug Design and Discovery*. Elsevier, 2014. DOI: 10.1016/c2011-0-07765-7.

[170] Stephen Neidle. "Quadruplex nucleic acids as targets for anticancer therapeutics." In: *Nature Reviews Chemistry* 1.5 (2017). DOI: 10.1038/s41570-017-0041.

[171] Stephanie M. Nelson, Lynnette R. Ferguson, and William A. Denny. "Non-covalent ligand/DNA interactions: Minor groove binding agents." In: *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 623.1-2 (2007), pp. 24–40. DOI: 10.1016/j.mrfmmm.2007.03.012.

[172] C. Nieuwland, F. Zaccaria, and C. Fonseca Guerra. "Understanding alkali metal cation affinities of multi-layer guanine quadruplex DNA." In: *Physical Chemistry Chemical Physics* 22.37 (2020), pp. 21108–21118. DOI: 10.1039/d0cp03433a.

[173] Henri J. Nussbaumer. "The Fast Fourier Transform." In: *Fast Fourier Transform and Convolution Algorithms*. Springer Berlin Heidelberg, 1981, pp. 80–111. DOI: 10.1007/978-3-662-00551-4_4.

[174] ONETEP:http://www.onetep.org/.

[175] OPENMX:http://www.openmx-square.org.

[176] Pablo Ordejon, E. Artacho, and J. M. Soler. "Mixed Approach to Incorporate Self-Consistency into Order-N LCAO Methods." In: *MRS Proceedings* 408 (1995), pp. 85–92. DOI: 10.1557/proc-408-85.

[177] Pablo Ordejón, Emilio Artacho, and José M. Soler. "Self-consistent order-Ndensity-functional calculations for very large systems." In: *Physical Review B* 53.16 (1996), R10441–R10444. DOI: 10.1103/physrevb.53.r10441.

[178] Taisuke Ozaki. "Efficient low-order scaling method for large-scale electronic structure calculations with localized basis functions." In: *Physical Review B* 82.7 (2010), p. 075131. DOI: 10.1103/physrevb.82.075131.

[179] P. J. de Pablo, F. Moreno-Herrero, J. Colchero, J. Gómez Herrero, P. Herrero, A. M. Baró, Pablo Ordejón, José M. Soler, and Emilio Artacho. "Absence of dc-Conductivity in-DNA." In: *Physical Review Letters* 85.23 (2000), pp. 4992–4995. DOI: 10.1103/physrevlett.85.4992.

[180] Gary N. Parkinson, Michael P. H. Lee, and Stephen Neidle. "Crystal structure of parallel quadruplexes from human telomeric DNA." In: *Nature* 417.6891 (2002), pp. 876–880. DOI: 10.1038/nature755.

[181] Robert G. Parr and Yang Weitao. *Density-Functional Theory of Atoms and Molecules*. OXFORD UNIV PR, 1994. 352 pp. ISBN: 0195092767. URL: https://www.ebook.de/de/product/3237600/robert_g_parr_yang_weitao_density_functional_theory_of_atoms_and_molecules.html.

[182] Robert Parr. *Density-functional theory of atoms and molecules*. New York Oxford England: Oxford University Press Clarendon Press, 1989. ISBN: 9780195042795.

[183] Anh Tuân Phan. "Human telomeric G-quadruplex: structures of DNA and RNA sequences." In: *FEBS Journal* 277.5 (2009), pp. 1107–1117. DOI: 10.1111/j.1742-4658.2009.07464.x.

[184] Anh Tuân Phan, Vitaly Kuryavyi, Kim Ngoc Luu, and Dinshaw J. Patel. "Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K solution †." In: *Nucleic Acids Research* 35.19 (2007), pp. 6517–6525. DOI: 10.1093/nar/gkm706.

[185] M. Piris and J. M. Ugalde. "Iterative diagonalization for orbital optimization in natural orbital functional theory." In: *Journal of Computational Chemistry* 30.13 (2009), pp. 2078–2086. DOI: 10.1002/jcc.21225.

[186] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley. "Electron correlation theories and their application to the study of simple reaction potential surfaces." In: *International Journal of Quantum Chemistry* 14.5 (1978), pp. 545–560. DOI: 10.1002/qua.560140503.

[187] Puja Prasad, Pijus K. Sasmal, Ritankar Majumdar, Rajan R. Dighe, and Akhil R. Chakravarty. "Photocytotoxicity and near-IR light DNA cleavage activity of oxovanadium(IV) Schiff base complexes having phenanthroline bases." In: *Inorganica Chimica Acta* 363.12 (2010), pp. 2743–2751. DOI: 10.1016/j.ica.2010.03.016.

[188] P. Pulay. "ImprovedSCF convergence acceleration." In: *Journal of Computational Chemistry* 3.4 (1982), pp. 556–560. DOI: 10.1002/jcc.540030413.

[189] Péter Pulay. "Convergence acceleration of iterative sequences. the case of scf iteration." In: *Chemical Physics Letters* 73.2 (1980), pp. 393–398. DOI: 10.1016/0009-2614(80)80396-4.

[190] Peter Pulay. "Localizability of dynamic electron correlation." In: *Chemical Physics Letters* 100.2 (1983), pp. 151–154. DOI: 10.1016/0009-2614(83)80703-9.

[191] Bathula Rajasekhar, Chandan Kumar, G. Premkumar, Mohd Aamir Bin Riyaz, P. T. V. Lakshmi, and Toka Swu. "Computational studies on G-quadruplex DNA-stabilizing property of novel Wittig-based Schiff-Base ligands and their copper(II) complexes." In: *Structural Chemistry* 30.3 (2018), pp. 727–742. DOI: 10.1007/s11224-018-1229-7.

[192] Andrew M. Rappe, Karin M. Rabe, Efthimios Kaxiras, and J. D. Joannopoulos. "Optimized pseudopotentials." In: *Physical Review B* 41.2 (1990), pp. 1227–1230. DOI: 10.1103/physrevb.41.1227.

[193] Guntram Rauhut and Hans-Joachim Werner. "Analytical energy gradients for local coupled-cluster methods." In: *Physical Chemistry Chemical Physics* 3.22 (2001), pp. 4853–4862. DOI: 10.1039/b105126c.

[194] David Řeha, Martin Kabeláč, Filip Ryjáček, Jiří Šponer, Judit E. Šponer, Marcus Elstner, Sándor Suhai, and Pavel Hobza. "Intercalators. 1. Nature of Stacking Interactions between Intercalators (Ethidium, Daunomycin, Ellipticine, and 4',6-Diaminide-2-phenylindole) and DNA Base Pairs. Ab Initio Quantum Chemical, Density Functional Theory, and Empirical Potential Study." In: *Journal of the American Chemical Society* 124.13 (2002), pp. 3366–3376. DOI: 10.1021/ja011490d.

[195] Roman V. Reshetnikov, Jiri Sponer, Olga I. Rassokhina, Alexei M. Kopylov, Philipp O. Tsvetkov, Alexander A. Makarov, and Andrey V. Golovin. "Cation binding to 15-TBA quadruplex DNA is a multiple-pathway cation-dependent process." In: *Nucleic Acids Research* 39.22 (2011), pp. 9789–9802. DOI: 10.1093/nar/gkr639.

[196] Jan Řezáč, Jindřich Fanfrlík, Dennis Salahub, and Pavel Hobza. "Semiempirical Quantum Chemical PM6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes." In: *Journal of Chemical Theory and Computation* 5.7 (2009), pp. 1749–1760. DOI: 10.1021/ct9000922.

[197] Jan Řezáč and Pavel Hobza. "A halogen-bonding correction for the semiempirical PM6 method." In: *Chemical Physics Letters* 506.4-6 (2011), pp. 286–289. DOI: 10.1016/j.cplett.2011.03.009.

[198] Jan Řezáč and Pavel Hobza. "Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods." In: *Journal of Chemical Theory and Computation* 8.1 (2011), pp. 141–151. DOI: 10.1021/ct200751e.

[199] Christoph Riplinger and Frank Neese. "An efficient and near linear scaling pair natural orbital based local coupled cluster method." In: *The Journal of Chemical Physics* 138.3 (2013), p. 034106. DOI: 10.1063/1.4773581.

[200] Christoph Riplinger, Barbara Sandhoefer, Andreas Hansen, and Frank Neese. "Natural triple excitations in local coupled cluster calculations with pair natural orbitals." In: *The Journal of Chemical Physics* 139.13 (2013), p. 134101. DOI: 10.1063/1.4821834.

[201] Raphaël Rodriguez, G. Dan Pantoş, Diana P. N. Gonçalves, Jeremy K. M. Sanders, and Shankar Balasubramanian. "Ligand-Driven G-Quadruplex Conformational Switching By Using an Unusual Mode of Interaction." In: *Angewandte Chemie International Edition* 46.28 (2007), pp. 5405–5407. DOI: 10.1002/anie.200605075.

[202] Barnett Rosenberg, Loretta Van Camp, and Thomas Krigas. "Inhibition of Cell Division in Escherichia coli by Electrolysis Products from a Platinum Electrode." In: *Nature* 205.4972 (1965), pp. 698–699. DOI: 10.1038/205698a0.

[203] Barnett Rosenberg, Loretta Vancamp, James E. Trosko, and Virginia H. Mansour. "Platinum Compounds: a New Class of Potent Antitumour Agents." In: *Nature* 222.5191 (1969), pp. 385–386. DOI: 10.1038/222385a0.

[204] F. Rosu, V. Gabelica, H. Poncelet, and E. De Pauw. "Tetramolecular G-quadruplex formation pathways studied by electrospray mass spectrometry." In: *Nucleic Acids Research* 38.15 (2010), pp. 5217–5225. DOI: 10.1093/nar/gkq208.

[205] Carmen L. Ruehl, Aaron H. M. Lim, Timothy Kench, David J. Mann, and Ramon Vilar. "An Octahedral Cobalt(III) Complex with Axial NH 3 Ligands that Templates and Selectively Stabilises G-quadruplex DNA." In: *Chemistry - A European Journal* 25.41 (2019), pp. 9691–9700. DOI: 10.1002/chem.201901569.

[206] Nicholas J. Russ and T. Daniel Crawford. "Local correlation in coupled cluster calculations of molecular response properties." In: *Chemical Physics Letters* 400.1-3 (2004), pp. 104–111. DOI: 10.1016/j.cplett.2004.10.083.

[207] SIESTA:https://departments.icmab.es/leem/siesta/.

[208] Svein Sæbø and Peter Pulay. "Local configuration interaction: An efficient approach for larger molecules." In: *Chemical Physics Letters* 113.1 (1985), pp. 13–18. DOI: 10.1016/0009-2614(85)85003-x.

[209] Svein Saebo and Peter Pulay. "The local correlation treatment. II. Implementation and tests." In: *The Journal of Chemical Physics* 88.3 (1988), pp. 1884–1890. DOI: 10.1063/1.454111.

[210] Masaaki Saitow, Ute Becker, Christoph Riplinger, Edward F. Valeev, and Frank Neese. "A new near-linear scaling, efficient and accurate, open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory." In: *The Journal of Chemical Physics* 146.16 (2017), p. 164105. DOI: 10.1063/1.4981521.

[211] Ángel Sánchez-González, Tarsila G. Castro, Manuel Melle-Franco, and Adrià Gil. "From groove binding to intercalation: unravelling the weak interactions and other factors modulating the modes of interaction between methylated phenanthroline-based drugs and duplex DNA." In: *Physical Chemistry Chemical Physics* 23.47 (2021), pp. 26680–26695. DOI: 10.1039/d1cp04529f.

[212] Ángel Sánchez-González and Adrià Gil. "Elucidating the intercalation of methylated 1,10-phenanthroline with DNA: the important weight of the CH/H interactions and the selectivity of CH/π and CH/n interactions." In: *RSC Advances* 11.3 (2021), pp. 1553–1563. DOI: 10.1039/d0ra07646e.

[213] Daniel Sánchez-Portal, Pablo Ordejón, Emilio Artacho, and José M. Soler. "Density-functional method for very large systems with LCAO basis sets." In: *International Journal of Quantum Chemistry* 65.5 (1997), pp. 453–461. DOI: 10.1002/(sici)1097-461x(1997)65:5<453::aid-qua9>3.0.co;2-v.

[214] Otto F. Sankey and David J. Niklewski. "Ab initiomulticenter tight-binding model for molecular-dynamics simulations and other applications in covalent systems." In: *Physical Review B* 40.6 (1989), pp. 3979–3995. DOI: 10.1103/physrevb.40.3979.

[215] Erwin Schrödinger. "Quantisierung als Eigenwertproblem." In: *Annalen der Physik* 384.4 (1926), pp. 361–376. DOI: 10.1002/andp.19263840404.

[216] Erwin Schrödinger. "The non-relativistic equation of the de Broglie waves." In: *Ann. der Phys* 79 (1926), pp. 361–489.

[217] Martin Schütz and Hans-Joachim Werner. "Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD)." In: *The Journal of Chemical Physics* 114.2 (2001), p. 661. DOI: 10.1063/1.1330207.

[218] Mark S. Searle, Allister J. Maynard, and Huw E. L. Williams. "DNA recognition by the anthracycline antibiotic respinomycin D: NMR structure of the intercalation complex with d(AGACGTCT)2." In: *Organic & Biomolecular Chemistry* 1.1 (2002), pp. 60–66. DOI: 10.1039/b208622k.

[219] Dipankar Sen and Walter Gilbert. "Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis." In: *Nature* 334.6180 (1988), pp. 364–366. DOI: 10.1038/334364a0.

[220] Hoi-Ling Seng, Sze-Tin Von, Kong-Wai Tan, Mohd Jamil Maah, Seik-Weng Ng, Raja Noor Zaliha Raja Abd Rahman, Ignez Caracelli, and Chew-Hee Ng. "Crystal structure, DNA binding studies, nucleolytic property and topoisomerase I inhibi-

tion of zinc complex with 1,10-phenanthroline and 3-methylpicolinic acid." In: *BioMetals* 23.1 (2009), pp. 99–118. DOI: 10.1007/s10534-009-9271-y.

[221] Isaiah Shavitt. "The history and evolution of configuration interaction." In: *Molecular Physics* 94.1 (1998), pp. 3–17. DOI: 10.1080/002689798168303.

[222] C. David Sherrill and Henry F. Schaefer. "The Configuration Interaction Method: Advances in Highly Correlated Approaches." In: *Advances in Quantum Chemistry*. Elsevier, 1999, pp. 143–269. DOI: 10.1016/s0065-3276(08)60532-8.

[223] Shuo Shi, Xiaoting Geng, Juan Zhao, Tianming Yao, Chaoran Wang, Danjing Yang, Lengfeng Zheng, and Liangnian Ji. "Interaction of [Ru(bpy)2(dppz)]2^{+} with human telomeric DNA: Preferential binding to G-quadruplexes over i-motif." In: *Biochimie* 92.4 (2010), pp. 370–377. DOI: 10.1016/j.biochi.2010.01.003.

[224] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, and L. H. Hurley. "Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription." In: *Proceedings of the National Academy of Sciences* 99.18 (2002), pp. 11593–11598. DOI: 10.1073/pnas.182256799.

[225] David J. Singh. *Planewaves, Pseudopotentials and the LAPW Method*. Springer US, 1994. DOI: 10.1007/978-1-4757-2312-0.

[226] Elisabeth Sjöstedt, Lars Nordström, and David Joseph Singh. "An alternative way of linearizing the augmented plane-wave method." In: *Solid State Communications* 114.1 (2000), pp. 15–20. DOI: 10.1016/s0038-1098(99)00577-3.

[227] Chris-Kriton Skylaris, Peter D. Haynes, Arash A. Mostofi, and Mike C. Payne. "IntroducingONETEP: Linear-scaling density functional simulations on parallel computers." In: *The Journal of Chemical Physics* 122.8 (2005), p. 084119. DOI: 10.1063/1.1839852.

[228] Nicolas Smargiasso, Frédéric Rosu, Wei Hsia, Pierre Colson, Erin Shammel Baker, Michael T. Bowers, Edwin De Pauw, and Valérie Gabelica. "G-Quadruplex DNA Assemblies: Loop Length, Cation Identity, and Multimer Formation†." In: *Journal*

of the American Chemical Society 130.31 (2008), pp. 10208–10216. DOI: 10.1021/ja801535e.

[229] José M Soler, Emilio Artacho, Julian D Gale, Alberto García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal. "The SIESTA method for ab initio order-N materials simulation." In: *Journal of Physics: Condensed Matter* 14.11 (2002), pp. 2745–2779. DOI: 10.1088/0953-8984/14/11/302.

[230] Petr Stadlbauer, Lukáš Trantírek, Thomas E. Cheatham, Jaroslav Koča, and Jiří Šponer. "Triplex intermediates in folding of human telomeric quadruplexes probed by microsecond-scale molecular dynamics simulations." In: *Biochimie* 105 (2014), pp. 22–35. DOI: 10.1016/j.biochi.2014.07.009.

[231] James J. P. Stewart. "Optimization of parameters for semiempirical methods I. Method." In: *Journal of Computational Chemistry* 10.2 (1989), pp. 209–220. DOI: 10.1002/jcc.540100208.

[232] James J. P. Stewart. "Optimization of parameters for semiempirical methods II. Applications." In: *Journal of Computational Chemistry* 10.2 (1989), pp. 221–264. DOI: 10.1002/jcc.540100209.

[233] James J. P. Stewart. "Optimization of parameters for semiempirical methods. III Extension of PM3 to Be, Mg, Zn, Ga, Ge, As, Se, Cd, In, Sn, Sb, Te, Hg, Tl, Pb, and Bi." In: *Journal of Computational Chemistry* 12.3 (1991), pp. 320–341. DOI: 10.1002/jcc.540120306.

[234] James J. P. Stewart. "Application of the PM6 method to modeling the solid state." In: *Journal of Molecular Modeling* 14.6 (2008), pp. 499–535. DOI: 10.1007/s00894-008-0299-7.

[235] James J. P. Stewart. "Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters." In: *Journal of Molecular Modeling* 19.1 (2012), pp. 1–32. DOI: 10.1007/s00894-012-1667-x.

[236] James Stewart. *Computational Chemistry MOPAC2016*. 2016.

[237] Anthony J. Stone. "Computation of charge-transfer energies by perturbation theory." In: *Chemical Physics Letters* 211.1 (1993), pp. 101–109. DOI: 10.1016/0009-2614(93)80058-w.

[238] Joseph E. Subotnik, Alex Sodt, and Martin Head-Gordon. "The limits of local correlation theory: Electronic delocalization and chemically smooth potential energy surfaces." In: *The Journal of Chemical Physics* 128.3 (2008), p. 034103. DOI: 10.1063/1.2821124.

[239] Jing Sun, Yan An, Li Zhang, Huo-Yan Chen, Yan Han, Yu-Jia Wang, Zong-Wan Mao, and Liang-Nian Ji. "Studies on synthesis, characterization, and G-quadruplex binding of Ru(II) complexes containing two dppz ligands." In: *Journal of Inorganic Biochemistry* 105.2 (2011), pp. 149–154. DOI: 10.1016/j.jinorgbio.2010.10.005.

[240] Alessio Terenzi, Riccardo Bonsignore, Angelo Spinello, Carla Gentile, Annamaria Martorana, Cosimo Ducani, Björn Högberg, Anna Maria Almerico, Antonino Lauria, and Giampaolo Barone. "Selective G-quadruplex stabilizers: Schiff-base metal complexes with anticancer activity." In: *RSC Adv.* 4.63 (2014), pp. 33245–33256. DOI: 10.1039/c4ra05355a.

[241] Alessio Terenzi, Daniela Lötsch, Sushilla van Schoonhoven, Alexander Roller, Christian R. Kowol, Walter Berger, Bernhard K. Keppler, and Giampaolo Barone. "Another step toward DNA selective targeting: NiII and CuII complexes of a Schiff base ligand able to bind gene promoter G-quadruplexes." In: *Dalton Transactions* 45.18 (2016), pp. 7758–7767. DOI: 10.1039/c6dt00648e.

[242] Walter Thiel and Alexander A. Voityuk. "Extension of the MNDO formalism tod orbitals: Integral approximations and preliminary numerical results." In: *Theoretica Chimica Acta* 81.6 (1992), pp. 391–404. DOI: 10.1007/bf01134863.

[243] Walter Thiel and Alexander A. Voityuk. "Extension ofMNDO tod orbitals: Parameters and results for the halogens." In: *International Journal of Quantum Chemistry* 44.5 (1992), pp. 807–829. DOI: 10.1002/qua.560440511.

[244] Walter Thiel and Alexander A. Voityuk. "Extension of MNDO to d Orbitals: Parameters and Results for the Second-Row Elements and for the Zinc Group." In: *The Journal of Physical Chemistry* 100.2 (1996), pp. 616–626. DOI: 10.1021/jp952148o.

[245] Eddy Thiriot and Gerald Monard. "Combining a genetic algorithm with a linear scaling semiempirical method for protein-ligand docking." In: *Journal of Molecular Structure: THEOCHEM* 898.1-3 (2009), pp. 31–41. DOI: 10.1016/j.theochem.2008.12.041.

[246] Julian Tirado-Rives and William L. Jorgensen. "Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules." In: *Journal of Chemical Theory and Computation* 4.2 (2008), pp. 297–306. DOI: 10.1021/ct700248k.

[247] J. Tomasi. "Thirty years of continuum solvation chemistry: a review, and prospects for the near future." In: *Theoretical Chemistry Accounts* 112.4 (2004). DOI: 10.1007/s00214-004-0582-3.

[248] N. Troullier and José Luriaas Martins. "Efficient pseudopotentials for plane-wave calculations. II. Operators for fast iterative diagonalization." In: *Physical Review B* 43.11 (1991), pp. 8861–8869. DOI: 10.1103/physrevb.43.8861.

[249] N. Troullier and José Luriaas Martins. "Efficient pseudopotentials for plane-wave calculations." In: *Physical Review B* 43.3 (1991), pp. 1993–2006. DOI: 10.1103/physrevb.43.1993.

[250] Vincenzo Tschinke and Tom Ziegler. "Gradient corrections to the Hartree-Fock-Slater exchange and their influence on bond energy calculations." In: *Theoretica Chimica Acta* 81.1-2 (1991), pp. 65–78. DOI: 10.1007/bf01113378.

[251] Eiji Tsuchida. "Augmented Orbital Minimization Method for Linear Scaling Electronic Structure Calculations." In: *Journal of the Physical Society of Japan* 76.3 (2007), p. 034708. DOI: 10.1143/jpsj.76.034708.

[252] Eiji Tsuchida and Masaru Tsukada. "Large-Scale Electronic-Structure Calculations Based on the Adaptive Finite-Element Method." In: *Journal of the Physical Society of Japan* 67.11 (1998), pp. 3844–3858. DOI: 10.1143/jpsj.67.3844.

[253] Seiji Tsuzuki and Asuka Fujii. "Nature and physical origin of CH/π interaction: significant difference from conventional hydrogen bonds." In: *Physical Chemistry Chemical Physics* 10.19 (2008), p. 2584. DOI: 10.1039/b718656h.

[254] David Vanderbilt. "Soft self-consistent pseudopotentials in a generalized eigenvalue formalism." In: *Physical Review B* 41.11 (1990), pp. 7892–7895. DOI: 10.1103/physrevb.41.7892.

[255] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler. "Chemistry with ADF." In: *Journal of Computational Chemistry* 22.9 (2001), pp. 931–967. DOI: 10.1002/jcc.1056.

[256] Florian Weigend. "Accurate Coulomb-fitting basis sets for H to Rn." In: *Physical Chemistry Chemical Physics* 8.9 (2006), p. 1057. DOI: 10.1039/b515623h.

[257] Florian Weigend and Reinhart Ahlrichs. "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy." In: *Physical Chemistry Chemical Physics* 7.18 (2005), p. 3297. DOI: 10.1039/b508541a.

[258] Hans-Joachim Werner and Peter J. Knowles. "An efficient internally contracted multiconfiguration-reference configuration interaction method." In: *The Journal of Chemical Physics* 89.9 (1988), pp. 5803–5814. DOI: 10.1063/1.455556.

[259] Hans-Joachim Werner and Klaus Pflüger. "Chapter 4. On the Selection of Domains and Orbital Pairs in Local Correlation Treatments." In: *Annual Reports in Computational Chemistry*. Elsevier, 2006, pp. 53–80. DOI: 10.1016/s1574-1400(06)02004-4.

[260] Lando P. Wolters, Nicole W. G. Smits, and Célia Fonseca Guerra. "Covalency in resonance-assisted halogen bonds demonstrated with cooperativity in N-halo-guanine quartets." In: *Physical Chemistry Chemical Physics* 17.3 (2015), pp. 1585–1592. DOI: 10.1039/c4cp03740e.

[261] Peng Wu, Dik-Lung Ma, Chung-Hang Leung, Siu-Cheong Yan, Nianyong Zhu, R. Abagyan, and Chi-Ming Che. "Stabilization of G-Quadruplex DNA with Platinum(II) Schiff Base Complexes: Luminescent Probe and Down-Regulation of c-mycOncogene Expression." In: *Chemistry - A European Journal* 15.47 (2009), pp. 13008–13021. DOI: 10.1002/chem.200901943.

[262] Yang Yang, Haibo Yu, Darrin York, Marcus Elstner, and Qiang Cui. "Description of Phosphate Hydrolysis Reactions with the Self-Consistent-Charge Density-Functional-Tight-Binding (SCC-DFTB) Theory. 1. Parameterization." In: *Journal of Chemical Theory and Computation* 4.12 (2008), pp. 2067–2084. DOI: 10.1021/ct800330d.

[263] Jun-Liang Yao, Xing Gao, Wenliang Sun, Shuo Shi, and Tian-Ming Yao. "[Ru(bpy)2dppz-idzo]$^{2+}$: a colorimetric molecular "light switch" and powerful stabilizer for G-quadruplex DNA." In: *Dalton Transactions* 42.16 (2013), p. 5661. DOI: 10.1039/c3dt32640c.

[264] Nusret Duygu Yilmazer, Pascal Heitel, Tobias Schwabe, and Martin Korth. "Benchmark of electronic structure methods for protein-ligand interactions based on high-level reference data." In: *Journal of Theoretical and Computational Chemistry* 14.01 (2015), p. 1540001. DOI: 10.1142/s0219633615400015.

[265] Nusret Duygu Yilmazer and Martin Korth. "Comparison of Molecular Mechanics, Semi-Empirical Quantum Mechanical, and Density Functional Theory Methods for Scoring Protein-Ligand Interactions." In: *The Journal of Physical Chemistry B* 117.27 (2013), pp. 8075–8084. DOI: 10.1021/jp402719k.

[266] F. Zaccaria, G. Paragi, and C. Fonseca Guerra. "The role of alkali metal cations in the stabilization of guanine quadruplexes: why K^+ is the best." In: *Physical Chemistry Chemical Physics* 18.31 (2016), pp. 20895–20904. DOI: 10.1039/c6cp01030j.

[267] Francesco Zaccaria and Célia Fonseca Guerra. "RNA versus DNA G-Quadruplex: The Origin of Increased Stability." In: *Chemistry - A European Journal* 24.61 (2018), pp. 16315–16322. DOI: 10.1002/chem.201803530.

[268] Marie Zgarbová, Jiří Šponer, Michal Otyepka, Thomas E. Cheatham, Rodrigo Galindo-Murillo, and Petr Jurečka. "Refinement of the Sugar-Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA." In: *Journal of Chemical Theory and Computation* 11.12 (2015), pp. 5723–5736. DOI: 10.1021/acs.jctc.5b00716.

[269] J. Zhang et al. "Recent Progress and Future Potential for Metal Complexes as Anticancer Drugs Targeting G-quadruplex DNA." In: *Current Medicinal Chemistry* 19.18 (2012), pp. 2957–2975. DOI: 10.2174/092986712800672067.

[270] Jun Zhou, Anne Bourdoncle, Frédéric Rosu, Valérie Gabelica, and Jean-Louis Mergny. "Tri-G-Quadruplex: Controlled Assembly of a G-Quadruplex Structure from Three G-Rich Strands." In: *Angewandte Chemie International Edition* 51.44 (2012), pp. 11002–11005. DOI: 10.1002/anie.201205390.

[271] Tom. Ziegler. "Approximate density functional theory as a practical tool in molecular energetics and dynamics." In: *Chemical Reviews* 91.5 (1991), pp. 651–667. DOI: 10.1021/cr00005a001.

[272] Jiří Šponer, Arnošt Mládek, Naďa Špačková, Xiaohui Cang, Thomas E. Cheatham, and Stefan Grimme. "Relative Stability of Different DNA Guanine Quadruplex Stem Topologies Derived Using Large-Scale Quantum-Chemical Computations." In: *Journal of the American Chemical Society* 135.26 (2013), pp. 9785–9796. DOI: 10.1021/ja402525c.