



Universitat de Girona

COMPUTATIONAL MODELING OF CHARGE TRANSFER IN NUCLEOBASE-AROMATIC AMINO ACID COMPLEXES

Cristina BUTCHOSA ROBLES

Dipòsit legal: GI. 1175-2012

<http://hdl.handle.net/10803/83529>



Computational Modeling of Charge Transfer in Nucleobase-Aromatic Amino Acid Complexes està subjecte a una llicència de [Reconeixement-NoComercial 3.0 No adaptada de Creative Commons](https://creativecommons.org/licenses/by-nc/3.0/)



PhD thesis:

Computational Modeling of Charge Transfer in Nucleobase-Aromatic Amino Acid Complexes

Cristina Butchosa Robles
2012

Programa de Doctorat en Ciències Experimentals i Sostenibilitat

PhD supervisors:

Dr. Sílvia Simon Rabasseda
Dr. Alexander A. Voityuk

Memòria presentada per a optar al títol de Doctora per la
Universitat de Girona



La professora Sílvia Simon Rabasseda a l'àrea de Química Física de la Universitat de Girona, i el professor Alexander A. Voityuk, investigador ICREA a l'Institut de Química Computacional de la Universitat de Girona,

CERTIFIQUEM:

Que aquest treball titulat "Computational Modeling of Charge Transfer in Nucleobase-Aromatic Amino Acid Complexes", que presenta la Cristina Butchosa Robles per a l'obtenció del títol de Doctora, ha estat realitzat sota la nostra direcció.

Signatura

Prof. Sílvia Simon Rabasseda

Prof. Alexander A. Voityuk

Girona, 10 de Maig de 2012

*a la Núria,
a tota la Família i els Amics
que m'han animat*

Summary

Oxidative damage on DNA produces radical cation states. These states are also called electron "holes". Electron holes can migrate long distances through the nucleobases stack, due to conductivity properties of DNA. Finally, the cationic charge could be trapped and most probably a mutagenic lesion will be initiated. However, if DNA interacts with a protein or peptide with higher affinity to cationic charges, the electron hole can be extracted from DNA by a charge transfer reaction. Although charge transfer reactions within DNA have been widely explored, not much is known about the charge transfer capabilities of DNA when it interacts with amino acids.

The present thesis models nucleobase-amino acid charge transfer reactions, which can serve as base for future computational investigations of charge transfer processes in DNA-protein systems. Guanine and adenine charge transfer reactions with aromatic amino acids (histidine, phenylalanine, tryptophan, and tyrosine) have been studied. Both π -stacked and T-shaped nucleobase-aromatic amino acid interactions have been found to produce fast charge transfer rates. Most of the aromatic amino acids are able to extract charges from DNA. Special attention has been paid to tryptophan because its redox properties. Tryptophan is the best aromatic amino acid to extract cationic charges from guanine and adenine.

The studied interactions have been shown to be extremely sensitive to conformational fluctuations. The electronic coupling between the moieties of the system, also the charge transfer rate, can change several orders of magnitude for small fluctuations of less than 1Å.

Nucleobase-Tryptophan interactions have been shown to be able to produce charge transfer reactions. However, their computational

modeling present some difficulties. Tryptophan has the first excited state close in energy to its ground state. For this, a 3-state treatment has been found to be necessary to describe its charge transfer reactions. A Multi-State level of theory should be used on the calculations involving tryptophan.

Resum

Les lesions oxidatives de l'ADN produeixen estats radical catió. Aquests estats també es poden anomenar "forats" d'electró. A causa de les propietats conductores de l'ADN, els forats d'electró poden migrar llargues distàncies a través de la seqüència de nucleobases. Finalment, la càrrega catiònica serà atrapada i, molt probablement, es produirà la iniciació d'una lesió mutagènica. No obstant, si l'ADN interacciona amb una proteïna o un pèptid que tingui una major afinitat per les càrregues catiòniques, el forat d'electró es pot extreure de l'ADN mitjançant una reacció de transferència de càrrega. Tot i que les reaccions de transferència de càrrega dins de l'ADN han estat extensament estudiades, les habilitats de transferència de càrrega quan aquest interacciona amb aminoàcids encara no ha estat investigades en profunditat.

Aquesta tesis doctoral modelitza reaccions de transferència de càrrega en sistemes compostos per una nucleobase i un aminoàcid. Un millor coneixement d'aquestes interaccions pot servir com a base per futures investigacions de processos de transferència de càrrega en sistemes ADN-proteïna.

S'han estudiat les reaccions de transferència de càrrega de la guanina o adenina amb aminoàcids aromàtics (histidina, fenilalanina, triptòfan i tirosina). Especialment, les interaccions amb el triptòfan, que gràcies a un potencial de ionització semblant al de la guanina pot estabilitzar millor que els altres aminoàcids un estat radical catió.

S'han considerat tant interaccions dels sistemes π com conformacions en forma T del les parelles nucleobase-aminoàcid, obtenint velocitats de reacció ràpides. La majoria d'aminoàcids aromàtics són capaços d'extreure càrregues de l'ADN.

Les interaccions estudiades s'han mostrat molt sensibles a fluctuacions conformacionals. L'acoblament electrònic entre les dues molècules del sistema, així com la velocitat de la transferència de càrrega, poden canviar diversos ordres de magnitud a causa de petites fluctuacions de menys d'1 Å.

Les interaccions Nucleobase-Triptòfan poden donar reaccions de transferència de càrrega entre les dues molècules. Tot i així, el seu modelatge computacional presenta algunes dificultats. El triptòfan té el primer estat excitat molt pròxim en energia al seu estat fonamental. Això fa que sigui necessària la utilització d'un tractament de 3-estats per descriure les seves reaccions de transferència de càrrega. Per tal de tenir una millor descripció del sistema s'ha d'utilitzar un nivell de càlcul basat en Multi-estats (CASPT2).

Acknowledgements

Bé, això s'acaba... i estic contenta, eh!!!... però també estic trista perquè fer el doctorat ha estat una etapa molt important de la meua vida. Aquests anys he après molt i he conegut molta gent fantàstica. En aquest agraïments segur que m'endescuidaré gent que es mereixeria ser-hi.

Quan vaig acabar la carrera tenia la sensació de no haver aprofundit en cap tema, per això la idea de fer un doctorat m'atreia força. Ara, fent balanç, crec que en se una mica de charge transfer i de química computacional... si ve, en això de la recerca sempre van apareixent nous interrogants, hi ha tantes coses que voldria investigar encara... Gràcies als meus directors de tesi, la Sílvia i l'Alexander i també als meus altres col·laboradors, dels quals he après moltíssim.

Vull agrair el suport que m'ha donat la meua família i especialment a la Lluïsa i la Tia. Estic segura que elles seran les més contentes si aconseguixo ser doctora. Núria, és una pena que siguis a Estocolm perquè no t'he pogut obligar ni una sola vegada a que et llegissis la tesi per ajudar-me a trobar-ne els errors. Que l'últim cop que vaig venir-te a veure vaig acabar comptant-te colònies de bacteris... com quan t'ajudava a fer treballs de plàstica de petita. Espero un cop de mà molt gran amb l'organització de les celebracions de la tesi (si tot va bé) que sinó el karma et passarà factura quan haguem de celebrar la teua... que segur que serà aviat!

La Núria i la Paula han estat el meu millor suport. Moltes gràcies per recordar-me que cal deixar l'ordinador de tant en tant i per el vostre assessorament gràfic. També agrair a en Yago el seu incalculable valor com a amic i com a traductor, i pels Mario Karts i els Dominions després de dinar.

A tots els amics de kendo, de rol (de taula i en viu), dels jocs de taula i del club de voleibol AVAP (tant a l'equip de patxanga com a les seniors supersexis de la secta Flow-Flow): moltes gràcies per la vostra paciència i per ajudar-me a distreure'm. Que la vida sense hobbies i sense vosaltres seria molt més trista. També a les noies del Vedruna (no, jo no vaig anar al Vedruna), els sopars amb vosaltres son estupendos!

Companys IQCians, moltes gràcies per tot!!! Vaig començar amb el treball experimental de la carrera al despatx 177 (si, no vaig estar mai al 166 que sóc jove jo) i vaig descobrir que encaixava millor a l'IQC que a cap altre grup o tribu social de la universitat. Deu ser que sóc una mica friki en el fons.... Però bé, de fet no encaixava perquè sigui friki (només) sinó perquè a l'IQC hi ha un ambient de treball immillorable!!! Les estones a Can Paco amb tota la colla de becaris quan vaig començar... mai he rigut tant. Els congressos, les excursions/peregrinacions de l'IQC, les JODETE, els moments als servei d'esports fent padels i futbols, els sopars i les tesis... aquests anys de doctorat estan plens de bons records. Llavors vam ser desterrats al Parc, i tot i que el despatx és més fashion, els companys de la facultat es trobaven a faltar. Als companys del despatx del Parc, em sap greu haver tingut la finestra tant tancada però ja sabeu que em constipò fàcilment... Carme, moltes gràcies per tot l'ajut i grans estones fent cafès.

A la Micra, encara que ella òbviament es una gata i no pot llegir. Algun dia rellegiré aquestes línies i recordaré la sensació de tornar a casa del Parc a la 1 de la nit i notar com puja al llit. El seu ronroneig sempre m'ha ajudat a dormir.

Finalment, gràcies a la Generalitat de Catalunya per els 3 anys de beca que m'han permès tot aquest periple d'estudiar un doctorat.

List of Figures

1.1	DNA as electronic wire.	1
1.2	Main steps of photoinduced CT in DNA.	3
1.3	Guanine structure.	6
1.4	Tryptophan (left) and Indole (right) structures.	6
1.5	Structure of Melatonin hormone.	7
1.6	DNA repair stimulated by CT	8
1.7	Schematized model of transcriptional activation of SoxR from a distance through DNA-mediated charge transport.	9
1.8	NCP	11
1.9	From left to right, the structures of A-, B- and Z-DNA.	12
1.10	DNA quadruplex	13
1.11	Van der Waals contacts between nucleobases and amino acids.	16
1.12	Composite nucleoside binding pockets produced by superposition of each base from 41 high resolution protein-RNA structures.	17
1.13	Parabolas of outer-sphere reorganization energy	19
1.14	Marcus parabolas for different redox reactions	20
1.15	Potential energy surfaces of initial (left parabola) and final (right parabola) states in Hole Transfer.	22
1.16	Fluctuations of the driving force (ΔG)	23
1.17	Fluctuations of conformational parameters of the base steps	25
2.1	For (G-Ind) $^{\bullet+}$ systems 3 states can be involved in the charge transfer reaction. One from Guanine (donor) and two from Indole (hole acceptor).	33
2.2	CAS and RAS MOs scheme.	45
2.3	Structural parameters that define the mutual arrange- ments of 2 moieties forming a complex.	50

7.1	IP of G, Ind, and A	78
7.2	π -stacked (G-Ind) $^{\bullet+}$ structures.	79
7.3	For (G-Ind) $^{\bullet+}$ systems, 3 states can be involved in the charge transfer reaction.	81
7.4	Representation of V^2 calculated using 3-state and the 2-state models.	85
7.5	V^2 calculated using 3-state model, for 6 different (G-Ind) $^{\bullet+}$ stacked conformations	86
7.6	Rate constant logarithm of two HT possible reactions, $k_{GS,CT}$ and $k_{LE,CT}$	87
8.1	π_1 stacked conformation and T3 (T-shaped) Guanine-Indole structures.	91
8.2	Comparison of V_{GS-CT}	94
8.3	Comparison of V_{LE-CT}	96
9.1	Structure of the guanine-tryptophan complexes.	100
9.2	Charge difference between the nucleobase and residue.	101
9.3	Dependence of the electronic coupling and free energy on the dimer conformation.	102
9.4	Conformational dependence of Charge, G-Trp	104
9.5	Dependence of ΔE and $\log V$ on G-Trp displacements on shift and slide parameters.	105
9.6	V^2 of G-Trp vs. A-Trp for twist rotations.	107
9.7	Dependence of ΔE and V on G-Trp rotations of tilt and roll.	108
9.8	Dependence of the conformational ΔE and V on displacement of G-Trp in the T-shaped complexes.	109
9.9	V^2 calculated with 3-state and the 2-state models.	110
9.10	Rate constant logarithm of two HT possible reactions, k_{GS-CT} and $k_{LE,CT}$	111
A.1	Nomenclature used for Guanine and Adenine faces.	122
A.2	Comparison of electronic couplings for A-X complexes calculated using the GMH and FCM schemes and the direct method (X-axis).	123

List of Tables

1.1	Rate constants for repair of DNA guanine radicals by aromatic amino acids	10
1.2	Van der Waals contacts between aromatic amino acids and nucleobases.	15
1.3	Distribution of stacking interactions in 41 high resolution protein-RNA structures.	17
7.1	MS-CASPT2 difference of adiabatic states energy and dipole moments.	83
7.2	λ_D values for the Cave's diagnostic method.	84
7.3	Ratio between the rates of HT reactions (between CT-GS or CT-LE states) calculated with the 2 or the 3-state methods.	88
8.1	Electronic coupling obtained for (G-Ind) $^{\bullet+}$ system between ground and charge transfer states, in meV. . . .	95
8.2	Electronic coupling obtained for (G-Ind) $^{\bullet+}$ system between locally excited and charge transfer states, in meV.	95
9.1	λ_D values for the Cave's diagnostic method.	109
A.1	Notation of dimer configurations in comparison to the used by Wetmore et al.	121
A.2	Comparison of the ET driving force ΔE and V obtained with different methodologies.	122
A.3	Results for complexes G-X , energies and couplings. . .	124
A.4	Results for complexes A-X , energies and couplings. . .	125
A.5	Atomic coordinates of G-Ind stacked reference structure in Angstroms.	126
A.6	Atomic coordinates of G-Ind T-shaped reference structure in Angstroms.	127

A.7	Atomic coordinates of A-Ind stacked reference structure in Angstroms.	128
A.8	Atomic coordinates of A-Ind T-shaped reference structure in Angstroms.	129
A.9	Atomic coordinates of $\pi 1$ structure in Angstroms (Total stack).	130
A.10	Atomic coordinates of $\pi 2$ structure in Angstroms (Shift 2 Å).	131
A.11	Atomic coordinates of $\pi 3$ structure in Angstroms (Slide -1.7Å).	132
A.12	Atomic coordinates of $\pi 4$ structure in Angstroms (Slide 2.7Å).	133
A.13	Atomic coordinates of $\pi 5$ structure in Angstroms (Twist 110°).	134
A.14	Atomic coordinates of $\pi 6$ structure in Angstroms (G-Ind) \bullet^+ structure where Ind is interacting with the other guanine side.	135
A.15	Electronic coupling, free energy and HT reaction rate of (G-Ind) \bullet^+ conformations. Results belongs both to the use of the 3-state (3s) and the 2-state (2s) GMH methods.	136
A.16	Atomic coordinates of T1 structure in Angstroms (Roll 100°).	137
A.17	Atomic coordinates of T2 structure in Angstroms (Tilt 50°).	138
A.18	Atomic coordinates of T3 structure in Angstroms.	139

List of Publications

The following publications have been included as chapters in this thesis:

1. Butchosa, C.; Simon, S.; Voityuk, A. A., Electron transfer from aromatic amino acids to guanine and adenine radical cations in π -stacked and T-shaped complexes. *Org. Biomol. Chem.* **2010**, 8, 1870-1875.
2. Butchosa, C.; Simon, S.; Voityuk, A. A., Conformational Dependence of the Electronic Coupling in Guanine-Tryptophan Complexes: A DFT Study. *Int. J. Quantum Chem.* **2012**, 112, 1838-1843.
3. Butchosa, C.; Simon, S.; Voityuk, A. A., Conformational dependence of the electronic coupling for hole transfer between adenine and tryptophan. *Comput. Theoret. Chem.* **2011**, 975, 38.

The material presented in chapter 7, "*Hole Transfer in Guanine-Indole Systems: A Multi-Configurational Study*", has been recently submitted with changes at Journal of Physical Chemistry as:

4. Butchosa, C.; Simon, S.; Blancafort, Ll.; Voityuk, A. A., MS-CASPT2 Study of Hole Transfer in Guanine-Indole complexes: Effective 2-State Treatment

Contents

Summary	vii
Resum	ix
Acknowledgements	xi
List of Figures	xiv
List of Tables	xvi
List of Publications	xvii
1 Introduction	1
1.1 CT in DNA-Protein systems	2
1.1.1 Charge transfer in DNA	2
1.1.2 DNA-protein interactions and CT	5
1.2 Struct. DNA-Protein	10
1.2.1 DNA structure	10
1.2.2 NB-aa interactions	14
1.3 Electron Transfer Theory	18
1.3.1 Charge transfer parameters	22
1.3.2 Excess Charge Delocalization	26
2 Methodology	29
2.1 Calculation of electronic coupling	30
2.1.1 Direct Method	30
2.1.2 2-state Generalized Mulliken-Hush	31
2.1.3 Fragment Charge Method	32
2.1.4 3-state GMH	32
2.1.5 Koopmans' Theorem Approximation	35
2.2 Density Functional Theory	37

2.2.1	Kohn-Sham equations	38
2.2.2	DFT Methods	39
2.3	Multi-Configurational Methods	42
2.3.1	CASSCF	43
2.4	Including dynamic electron correlation	46
2.4.1	CASPT2	47
2.4.2	MS-CASPT2	49
2.5	Structures	50
2.5.1	Structural parameters	50
3	Objectives	51
4	ET in G-X and A-X systems	53
5	Dependence of V in G-Trp	61
6	Dependence of V in A-Trp	69
7	HT in G-Ind systems: MC study	75
7.1	Abstract	76
7.2	Introduction	76
7.3	Computational details	77
7.3.1	Geometries	78
7.4	Methodology	79
7.4.1	2-state GMH method	80
7.4.2	3-state GMH method	81
7.5	Results and discussion	82
7.5.1	Dipole moments and energies	82
7.5.2	Electronic couplings	84
7.5.3	HT rates with 3-state method	85
7.5.4	Comparison $k_{2\text{-state}}/k_{3\text{-state}}$	87
7.6	Conclusions	88
8	HT in G-Ind; DFT vs. MS-CASPT2	89
8.1	Abstract	90
8.2	Introduction	90
8.3	Computational details	91
8.3.1	LC-DFT theoretical details	92
8.4	Results and discussion	93
8.4.1	$V_{\text{GS-CT}}$	93
8.4.2	$V_{\text{LE-CT}}$	96

8.5	Conclusions	97
9	Results and discussion	99
9.1	ET in G-X and A-X systems	100
9.1.1	G-X and A-X distribution of charge	101
9.1.2	G-X and A-X V and ΔG	102
9.1.3	G-X and A-X HT rates	103
9.2	Conform. dependence of V in G-Trp and A-Trp	103
9.2.1	Electronic couplings of π -stacked conformations	105
9.2.2	Electronic couplings of T-shaped G-Trp and A-Trp conformations	106
9.3	HT in G-Ind; MS-Configurational Study	107
9.3.1	Electronic couplings of G-Ind systems	110
9.3.2	Hole transfer rates using the 3-state method	111
9.4	Results DFT vs. MS-CASPT2	112
9.4.1	$V_{\text{GS-CT}}$	112
9.4.2	$V_{\text{LE-CT}}$	112
10	Conclusions	115
	List of Acronyms	119
A	Appendix	121
A.1	SM chapter 4	121
A.2	SM chapter 5	126
A.3	SM chapter 6	128
A.4	SM chapter 7	130
A.5	SM chapter 8	137
	Bibliography	141

Chapter 1

Introduction

The double-helix structure of DNA is composed by nucleobase pairs and sugar backbones. It is well known from the biological point of view, as it contains the genetic instructions used in the development and functioning of all known living organisms (with the exception of RNA viruses). From a more chemical-physical context, this structure presents electronic properties that make DNA an electron organic conductor. Nine years after the discovery of DNA's structure by Watson and Crick,¹ Eley and Spivey² were the first to suggest that π - π interactions between stacked base pairs in double-strand DNA could act as electronic wires.



Figure 1.1: DNA as electronic wire.

Recently, different experimental measurements on single DNA molecules have been conducted. For example, Barton's group measured the fluorescence produced by an excited molecule and found that it could not emit light when attached to a DNA molecule. This suggests that the charge on the excited donor molecule passes to a nearby acceptor molecule in the "wire" of DNA causing this fluorescence quenching. They experimentally proved fast photoinduced electron transfer up to distances of over 40 Å.³⁻⁵

In recent years, much progress was done in the study of charge transport on DNA, both on experimental and theoretical areas, and now the main processes of Charge Transfer (CT) are understood. The migration of charge (radical or cation) can take place over distances up to few nanometers.

Several processes and applications use the electronic properties of DNA. For example, DNA damage and repair⁶ as well as the design of nanoelectronic devices or circuits in the nanotechnology field of material science.^{7,8}

Regarding DNA repair, oxidative stress can damage DNA because it can form an intermediate radical cation of guanine. Long-distance Hole Transfer (HT) could play a vital role to protect DNA from these oxidative damages. The charge from DNA can be transferred to an interacting molecule.⁹

1.1 Charge Transfer in DNA and DNA-Protein systems

1.1.1 Charge transfer in DNA

The migration of charge in DNA can occur both in reduced and oxidized DNA. In reduced DNA, it is an electron that migrates along DNA. This process is called Excess Electron Transfer (EET). In the case of oxidized DNA an electron hole, a radical cation state, is transferred. This process is known as Hole Transfer (HT). Since guanine has the lowest oxidation potential of the 4 nucleobases, Hole Transfer involves the formation of a guanine radical cation intermediate ($G^{\bullet+}$). Most experimental data have been obtained for hole transport processes.

The currently accepted method for HT reactions in DNA is a combination of G-hopping and super-exchange.⁶ Figure 1.2 shows the main steps of these reactions. Experimentally, the insertion of a chromophore to initiate the HT reaction is needed.^{10,11} A chromophore (d) intercalated in the DNA strand is irradiated. This produces the injection of a radical cation in the sequence of bases into an adjacent guanine (G_1). Another guanine (G_2) adjacent to $G_1^{\bullet+}$ can donate

an electron to it producing the movement of the charge hole from $G_1^{\bullet+}$ to G_2 , and getting G_1 and $G_2^{\bullet+}$. The charge jumps from one guanine to the next (G-hopping). These G-hopping steps are mediated by super-exchange.

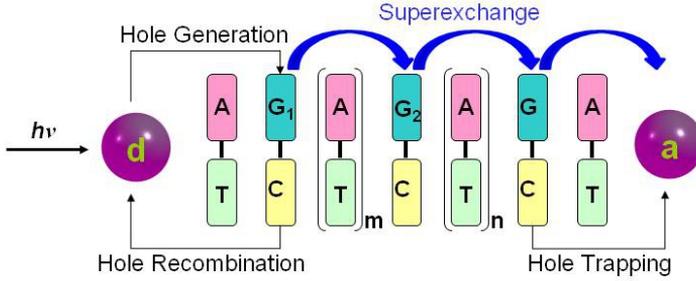


Figure 1.2: Main steps of photoinduced CT in DNA.

The radical cation $G^{\bullet+}$ migrates through DNA until it reacts with a water molecule. The efficiency of this process strongly depends on the sequence of base pairs. It has been experimentally shown that HT may occur over distances of 200\AA .¹²⁻¹⁴ Although $(AT)_n$ small bridges are not usually oxidized during the hole transfer process, for some large bridges A-hopping is also possible.^{15,16}

The rate of this HT mechanism depends exponentially on the distance between the donor and acceptor sites:

$$k_{da} = k_0 \exp(-\beta R_{da}) \quad (1.1)$$

β stands for the falloff parameter that determines the conductivity of the bridge. Usually β values range from 0.6 to 1.4\AA^{-1} . For $\beta=0.7 \text{\AA}^{-1}$, the value of k_{DA} decreases by an order of magnitude for each AT pair to be added between the donor and acceptor in DNA strand. If $\beta=1.4 \text{\AA}^{-1}$, the reduction of k_{DA} is two orders of magnitude.

As previously said, the super-exchange mechanism seems to be suitable for short distances below 15\AA . To describe long-distance charge transfer a multistep hopping mechanism is more adequate. The distance between the donor and the acceptor is divided into several short tunneling steps. Super-exchange controls the probability of these

steps. The hopping rate is proportional to $N^{-\eta}$, where N is the number of steps, and η is a parameter that takes values between 1 and 2.^{1,2}

A variety of experimental data has shown consistency with this multistep hopping mechanism.⁶ However, experiments suggest that, for $G(A)_nG$ sequences where n is bigger than three, $G^{\bullet+}$ oxidizes the next adenine and then A-hopping through A's is possible.¹⁷ Moreover, in a previous work by Voityuk *et al.* it has been shown that fluctuations in DNA environment can induce the hole transfer from $G^{\bullet+}$ to A.¹⁶ Surrounding water molecules can also affect the HT process, the charge transfer side reaction from $G^{\bullet+}$ to a water molecule is possible. However, it does not affect much the long range HT in DNA reaction.¹⁸ The solvent effect on DNA and DNA-protein charge transfer reactions mainly influences the reorganization energy of the system and their energetics.¹⁹ Also structural constraints could appear due to the presence of water molecules.²⁰ The charge transfer mechanism through DNA has a sensitive dependence on the complex structure and dynamics of DNA and the interaction with the solvent. The fluctuation of counterions, strongly counterbalanced by the surrounding water, leads to large oscillations of onsite energies, which govern the energetics of hole propagation along the DNA strand. Elstner's¹⁹ work concludes that the electronic couplings depend only on DNA conformation and are not affected by the solvent. In particular, the onsite energies are strongly correlated between neighboring nucleobases, indicating that a conformational-gating type of mechanism may be induced by the collective environmental degrees of freedom.

Although this work is focused on the hole transport (HT), the excess electron transport (EET) is also noteworthy. In these EET reactions, it is a radical-anion which is transported through DNA strand. Of the four DNA bases, thymine and cytosine are more easily reducible than purine nucleobases. Thus, pyrimidine bases are able to carry electron excess charge. Thymine has the strongest electron affinity, this suggests that excess charge could probably hop via thymine basis. Nowadays, in computational chemistry, EET through DNA is a rapidly growing topic both from mechanistic and dynamic points of view.^{21,22} It has also attracted much experimental interest, with some photochemical and spectroscopic studies published recently.²³⁻²⁵

1.1.2 DNA-protein interactions and Charge Transfer

This thesis is focused on the characterization of charge transfer reactions between aromatic amino acids (principally tryptophan) and adenine or guanine nucleobases, in order to model charge transfer reactions on DNA-protein interaction.

The stabilization of hole states on guanine bases is strongly dependent on the electrostatic interaction between NB and amino acids,²⁶ and the hole transport can be terminated because of these interactions. Furthermore, an excess charge can move from the π stack to an amino acid residue causing essential changes in hydrogen bonds and the electrostatic interaction between the NB and polypeptides followed by proton transfer and conformational transformation.

Because of the nucleobase-protein interactions (NB-protein), the structural and electronic parameters of NB-polypeptide complexes may considerably deviate from those of DNA. In particular, underwound and overwound DNA conformations are formed. Such structural changes affect CT properties considerably. Thus, by measuring the efficiency of CT within a NB-amino acid complex one can derive useful information about NB-protein interactions.

Guanine-Tryptophan Charge Transfer

Guanine-Tryptophan interaction seems to be the best pair between all the possible combinations of nucleobases and aromatic amino acids to perform HT reactions. What makes this G-Ind interaction so optimal to perform HT reactions from DNA to a protein or peptide? Both molecules present good hole acceptor properties.

- Guanine has the lowest ionization potential (8.02 eV) between the 4 nucleobases. Thus, cationic charges are more stabilized in this base. Guanine cationic radicals ($G^{\cdot+}$) are the most probable result of a DNA one-electron oxidation.²⁷⁻²⁹
- Tryptophan aromatic amino acid is well known for its role as hole acceptor on hole transfer reactions.^{30,31} Moreover its experimental ionization potential 7.90 eV is very similar to guanine's 8.02 eV.²⁹ One example of their importance in biological processes is its role on the photoactivation of photolyase protein.³² Photolyase uses reduced flavin cofactor ($FADH^-$) to

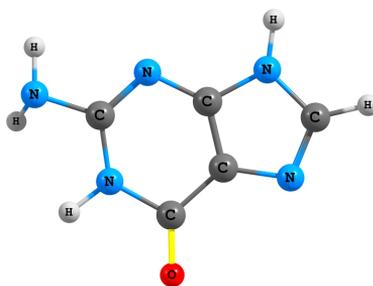


Figure 1.3: Guanine structure.

repair cyclobutane pyrimidine dimers (CPDs) which are the main UV-induced lesions on DNA.^{33–37} There is a chain of three Trps, well conserved through all known photolyases, which photoactivate the FADH^0 reducing it mediating an electron transfer reaction which takes 30 ps by a hopping mechanism.³⁸

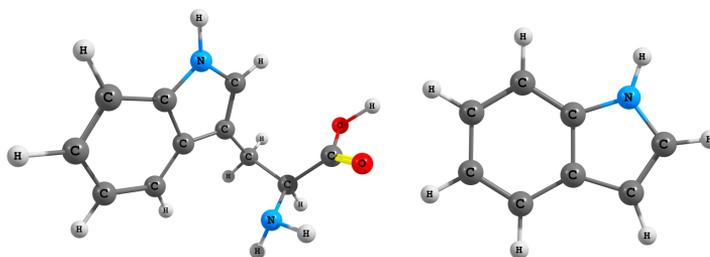


Figure 1.4: Tryptophan (left) and Indole (right) structures.

Moreover, the indole (Ind) moiety of tryptophan (see Figure 1.4) is known to be a successful antioxidant acting as free radical scavenger and broad-spectrum antioxidant in living organisms. One example of this antioxidant activity is the melatonin (MLT) hormone (see its structure on Figure 1.5), produced by the pineal gland from tryptophan. MLT does not only act as scavenger from free radicals but also stabilizes the membrane cells making them more resistant to oxidative damage. Furthermore, MLT can cross all the morphophysiological barriers and can protect from lipids in the cell membrane to DNA in the nucleus. Experimental efforts have been done on synthesizing MLT analogues, which

can improve MLT therapeutic capabilities.



Figure 1.5: Structure of Melatonin hormone.

Thus, all these reasons make the G-Trp pair the most suitable system within Nucleobase-Aromatic Amino Acid (NB-*aaa*) ones to produce a HT reaction.

Previous Charge Transfer studies on DNA-protein systems

Long-range oxidative damage to DNA has been shown to be extremely sensitive to protein-induced DNA distortions.³⁹ DNA-protein interactions can modulate electron-transfer chemistry in DNA.

Experimental studies have shown that oxidative stress, in the form of reactive oxygen species or other oxidative attacks, threatens cell survival, and is implicated in DNA damage, aging, and cancer. Barton and co-workers explored the possible role of DNA CT in DNA repair. First they demonstrated that redox activity required DNA binding for MutY,⁴⁰ a base excision repair (BER) enzyme from *Escherichia coli* that acts as a glycosylase to remove adenine from G:A and 7,8-dihydro-8-oxo-2-deoxyguanosine:A mismatches. Following this line of investigation, an interesting work about CT in DNA-protein systems carried out by Yavin et al.,⁴¹ on Barton's group is: "*Protein-DNA charge transport: Redox activation of a DNA repair protein by guanine radical*". It proposes a detection strategy for base excision repair (BER) enzymes. It uses DNA-mediated CT stimulated by guanine radicals (Figure 1.6).

The guanine radicals, formed under oxidative stress, are reduced and hence repaired through DNA-mediated electron transfer from a BER enzyme. Oxidation of the repair protein drives CT to an

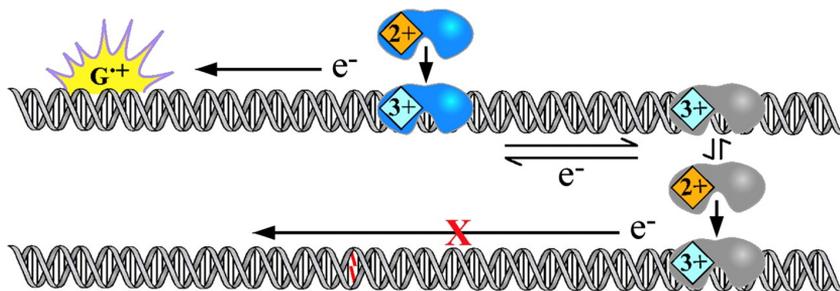


Figure 1.6: Model for detection strategy for BER enzymes using DNA-mediated CT stimulated by guanine radicals.⁴¹

alternate repair protein bound at a distant site, thereby promoting the redistribution of DNA repair proteins on genomic sites. On Yavin's work,⁴¹ the data describe MutY oxidation, but other BER enzymes containing $[4\text{Fe-4S}]^{2+}$ clusters show equivalent DNA-bound redox potentials. In solution BER enzymes are hardly oxidized because they have a $[4\text{Fe-4S}]^{2+}$ cluster. However, a BER-DNA binding shifts the cluster potential, promoting its oxidation to $[4\text{Fe-4S}]^{3+}$, with DNA-mediated CT to another oxidized repair protein bound at a distant site of the duplex. The reduction of this distant DNA-bound repair protein facilitates its dissociation from DNA and the relocation onto another site.

DNA lesions are signaled because when the protein binds to a region nearby a DNA lesion, DNA-mediated CT cannot occur. Then the repair protein progressively moves on a slower time scale to the site of the lesion and carries out its repair. In this way, DNA CT provides a route to redistribute the repair proteins onto regions of the genome containing DNA lesions.

Barton and co-workers have demonstrated that cells, more specifically DNA-protein interactions, are able to detect and respond to conditions of oxidative stress. They show that it is possible, thanks to the conductive properties of DNA, signal the information of oxidative stress across the genome till a distant protein that can extract the charge (DNA-mediated redox signaling for transcriptional activation of SoxR).⁴²

Experimental studies also pointed out that DNA interactions with

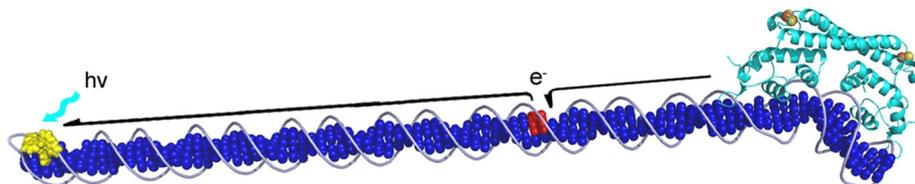


Figure 1.7: Schematized model of transcriptional activation of SoxR from a distance through DNA-mediated charge transport. Here, a tethered metal complex (yellow) is used to inject an electron hole into DNA base pair stack (dark blue) so as to generate a guanine radical (red). DNA-mediated charge transport from SoxR (light blue), bound at its promoter site, to the guanine radical fills the hole and leads to oxidation and activation of SoxR.⁴²

peptides can be useful to repair the oxidative damage of DNA. Some amino acids can reduce guanyl radicals formed when DNA is oxidized. Milligan experiments⁴³ (see Table 1.1) show that tyrosine could play a protective role mediated by charge transfer reactions. However, Tryptophan (Trp) is being the most reactive amino acid. The ionization potential of Trp amino acid (7.90 eV) is practically the same as guanine (8.02 eV) and their aromatic properties make this amino acid the most suitable to reduce oxidized DNA by a charge transfer reaction.

The theoretical study "*Hole Transfer Energetics in Structurally Distorted DNA: The Nucleosome Core Particle*", carried out by Voityuk and Davis²⁶ pointed that DNA distortion in Nucleosome Core particle (NCP) can drastically modify CT reactions between DNA nucleobases (Figure 1.8).

In Voityuk's work, quantum-chemical calculations (NDDO-G semiempirical theory) have been employed to study how the electrostatic interactions between DNA nucleobases and the surrounding protein and water molecules, as well as structural changes in DNA arising from compaction into a NCP, modify the energetics of hole transfer reaction between guanine sites.

Table 1.1: Rate constants for repair of DNA guanine radicals by amino acids, simple amino acids derivatives, and selected structurally similar compounds.⁴³

Structure	Derivative	$k = \text{dm}^3 \text{mol}^{-1} \text{s}^{-1}$
tryptophan (indole)	tryptophan	1×10^7
	Trp-Gly	7.7×10^6
	tryptophanamide	4.5×10^7
	Lys-Trp-Lys	9.2×10^7
tyrosine (phenol)	4-hydroxyphenylacetic acid	3×10^5
	3-(4'-hydroxyphenyl)propionic acid	2.8×10^5
	tyrosine	4×10^5
	N-acetyltyrosinamide	8.3×10^5
	tyramine	4.4×10^6
	tyrosinamide	1×10^6
	Lys-Tyr-Lys	3.6×10^6

Results show that structural distortions of DNA can have dramatic consequences for the stability of a guanine radical cation. These effects must be taken into account while modeling of in vivo DNA CT and in the interpretation of experimental findings. Considering the electrostatic potential arising from the water and basic histone proteins, DNA-histone contacts (particularly between arginine residues and the DNA minor groove) destabilize the cationic state on specific guanine residues. Thus, NB contacts with basic amino acids can change DNA sites preferred for hole stability.

Both theoretical and experimental studies have demonstrated that DNA-protein interactions can drastically affect electronic hole migration inside oxidized DNA. Moreover, some aromatic amino acids, such as Tryptophan, can repair DNA oxidative damage because they are able to reduce guanine radical cations. Charge transfer reactions are possible in DNA-protein interactions.

1.2 Structural aspects of DNA-Protein complexes

1.2.1 DNA structure

DNA consists of two long polymers made of simple units called nucleotides. The backbones of these nucleotides are made of sugars and phosphate groups. These backbones are disposed in two strands,

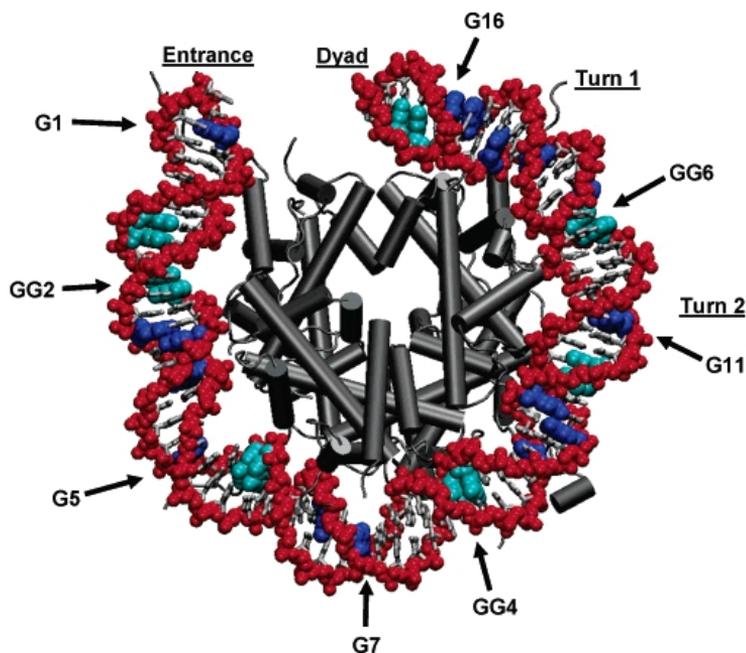


Figure 1.8: View of the nucleosome core particle rendered using the coordinates from PDB code 1KX5. The histone octamer is in gray, and only one-half of the palindromic DNA structure is shown. The DNA backbone is rendered in red, and the nucleobases are in gray. All 30 guanines are colored, with those in single G steps in dark blue and those in GG tracts in cyan. The locations of representative guanine residues are indicated.²⁶

which run in opposite directions. The other moiety of nucleotides are nucleobases. There are four types of nucleobases: Guanine, Adenine, Cytosine, and Thymine. These bases are coupled by hydrogen bonds interactions, Guanine with Cytosine and Adenine with Thymine. The structure a DNA molecule depends on its environment. In aqueous environments, including the majority of DNA in a cell, B-DNA is the most common structure (Figure 1.9). The A-DNA structure is dominates in dehydrated samples and is similar to the double-stranded RNA and DNA/RNA hybrids. Z-DNA is a rarer structure found in DNA bound to certain proteins. As it is shown in Figure 1.9, the relative position of these base pairs is almost parallel.

These nucleobases are bonded via sugar to negatively charged phosphate group. To neutralize these negative charges, DNA is surrounded

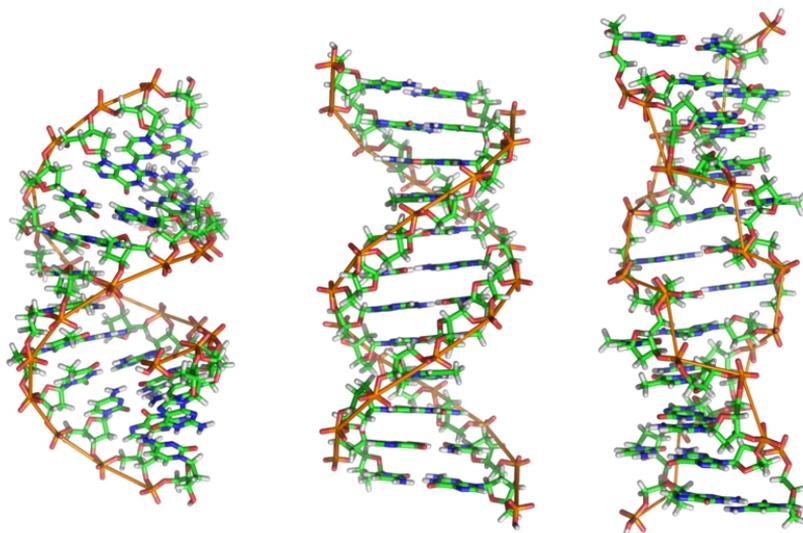


Figure 1.9: From left to right, the structures of A-, B- and Z-DNA.

by positive counterions like ions of Na^+ . These stabilizing entities may affect the charge transfer process. There are fluctuations in DNA structure that affect the charge-transfer rate. To calculate the charge transfer in DNA, both DNA and medium dynamics should be taken into account. Each different sequence of base pairs has its specific CT rate.

DNA is quite resistant to photoexcitation because the excited singlet state of nucleobases has a lifetime lower than 1 ps. This renders that to conduct CT experiments, the insertion or bonding of a chromophore in DNA strand is necessary to produce the charge transfer initiation by photoexcitation. Different studies have shown that the efficiency of hole transfer depends on two key factors: the gap between the donor and the bridge electronic levels, and the coupling between the corresponding orbitals.^{44,45} However, the insertion in DNA strand of a chromophore, acting as a donor site, can change significantly DNA structure. This can also affect the charge transfer rate. Another factor that may influence the charge transfer is the migration of charge through DNA strand. It will change the pKa values of the surrounding nucleobases favoring proton transfer reactions, and resulting in a concerted processes called Proton-Coupled Electron Transfer, which can affect a lot the charge transfer reaction.^{46,47}

DNA Quadruplex conformation

DNA is usually associated to a double-helix stacking of nucleobases pairs. However other conformations are also possible. For example, in some guanine-rich regions of DNA, such as telomeres or gene precursors, four-stranded quadruplex conformations (G-quadruplex) can be formed.⁴⁸⁻⁵⁰

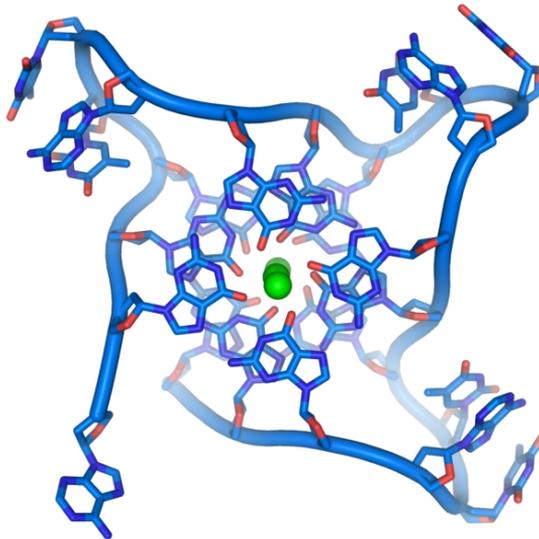


Figure 1.10: Crystal structure of parallel quadruplexes from human telomeric DNA. The DNA strand (blue) circles the bases that stack together in the center around three co-ordinated metal ions (green). Produced from NDB ID: UD0017 by Thomas Splettstoesser

A general G-quadruplex forming sequence can be defined as:



Where N_{L1-3} are loops that could vary in length and sequence. Guanine nucleobases are in the center of the G-quadruplex while the loops are positioned in the exterior helping to stabilize the system. The spaces between loops are limited by the charge of phosphodiester backbones.⁵¹ The core of G-quadruplex can be formed in several orientations (parallel or anti-parallel), and also the loops themselves are able to have elements of secondary structure. Moreover, recently some guanine nucleobases inside the loops have been found that can also be considered as a part of G-quadruplex core.⁵²⁻⁵⁴ This fact adds

complexity to a G-quadruplex structure prediction. Some structures have already been observed experimentally by NMR or crystallography techniques. They could serve as references starting points to structural predictions using computational methods.^{55,56}

The physicochemical properties of these G-quadruplex structures strongly differ from the ones of standard double-stranded DNA (dsDNA). Thus, designing molecules with higher affinity to bind G-quadruplex than dsDNA is feasible.^{57,58}

G-quadruplex conformations (especially for strands laying in parallel orientation) have been proved to be thermally more stable than dsDNA. Moreover it has a larger polarizability and a higher conductance. The improved charge transfer (CT) properties of G-quadruplex are attributed to the increase of structural stability and the higher overlapping of the π -orbitals of the system.⁵⁹ G-quadruplex DNA shows excellent characteristics to be employed in nano-devices.^{60,61} However the deviation of behavior affects CT processes inside DNA. Moreover these systems have been shown to be extremely sensitive to oxidative damage. For example telomeric DNA is harder to repair from oxidative damage than elsewhere in the chromosome. Oxidative stress of telomeres accelerates telomere loss and hastens cell senescence. However some studies postulate that G-quadruplex could act as a sink to cationic charges protecting more sensitive regions of the DNA.^{62,63}

Modeling DNA

DNA structure is highly complex with many atoms in its structure. Different options can be used to study these kind of systems. The first option is to include the maximum number of atoms performing low level computational calculations. The second option is to reduce the studied system to its most significant molecules and atoms, and perform high level calculations on it. The fragment selected for these calculations must be relevant for the studied reaction or interaction. DNA and their environment constantly suffer thermal fluctuations which can be addressed with molecular dynamics.

1.2.2 Architecture of Nucleobase-amino acid interactions

Several studies employed X-ray structures of DNA-protein and RNA-protein interactions from PDB, in order to analyze the probability of

nucleobase-amino acid contacts and their nature.^{64,65}

DNA-protein interactions

Focused on the understanding of nucleobase-amino acid recognition, the paper of Luscombe *et al.*,⁶⁴ provide information about nucleobases-amino acid contacts in a set of 129 DNA-protein structures. The studied contacts in this paper were hydrogen bonds, water mediated bonds and Van der Waals (VdW) interactions, the last one is the most relevant in charge transfer studies, specially the ones between guanine or adenine with aromatic amino acids. In Table 1.2, the frequencies of NB-aromatic amino acid Van der Waals contacts are listed. As can be seen, the expected contacts based on random interactions between the NB and aromatic amino acids (values in parenthesis in Table 1.2) can substantially differ from the real ones. Considering the whole set of studied contacts in the paper by Luscombe,⁶⁴ 2/3 of the total DNA-protein interactions analyzed comprise van der Waals contacts. Thus, these kind of contacts are of great importance in these systems. A graphical representation of these NB-amino acid Van der Waals contacts can be seen in Figure 1.11.

Table 1.2: Van der Waals contacts between aromatic amino acids and nucleobases. In parenthesis, the expected number of contacts supposing a random protein-DNA docking.⁶⁴

	Cytosine	Guanine	Adenine	Uracil
Phenylalanine	22 (3.4)	7 (3.5)	22 (4.1)	29 (6.9)
Histidine	9 (4.0)	12 (4.2)	- (4.9)	1 (8.2)
Tyrosine	1 (5.2)	6 (5.5)	- (6.4)	9 (10.8)
Tryptophan	8 (2.2)	- (2.1)	1 (2.5)	5 (4.2)

G-Trp and A-Trp contacts are not frequent in DNA-protein interactions. However, in the following chapters of this thesis they become of much more relevance because, as it was explained before, tryptophan has the best ionization potential among the four aromatic amino acids to allow charge transfer reactions.

RNA-protein interactions

The work of Morozova *et al.*⁶⁵ is focused on the characterization of the binding pockets formed by 41 protein-RNA complexes. Studying

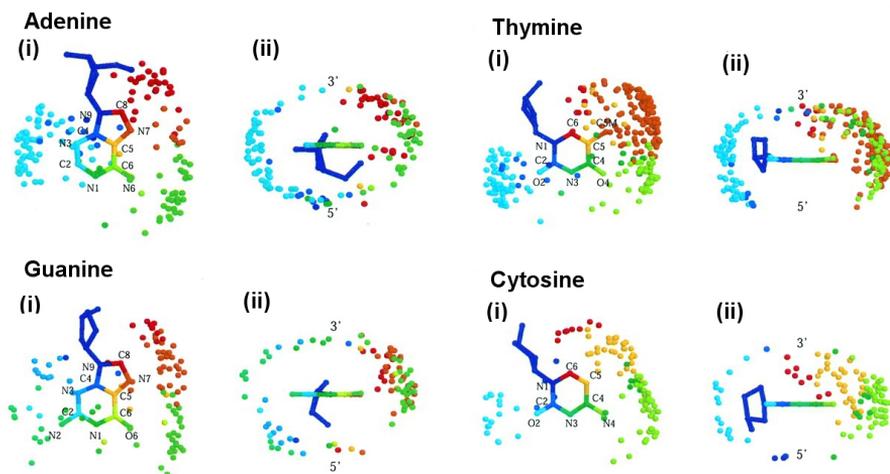


Figure 1.11: Van der Waals contacts between nucleobases and amino acids. Protein and DNA atoms interacting are identified with the same color. Two orientations are shown: i) interactions above the plane of the bases (3'-end) and ii) on the base-pairing edges.⁶⁴

these interactions one can derive how the nucleobases are recognized by the different proteins. RNA-protein interactions present much more variety of geometries than DNA-protein ones. This is due to the decrease of steric impediments derived from the unpaired nucleobases. Thus, protein-RNA interactions can result in the formation of a highly specific binding pocket around RNA bases, but the size, shape and non-polar binding patterns differ between specific RNA bases. The differences between the nucleobases allow proteins to make specific interactions with only few contacts, such as a pair of H-bonds in some cases as can be seen in Figure 1.12.

The stacking between aromatic amino acids and nucleobases can maximize a charge transfer reaction. In the study of Morozova,⁶⁵ over 41 protein-RNA interactions, nucleobase-amino acid stacked contacts are not very frequent, see Table 1.3.

As it was pointed out previously, guanine-aromatic amino acid interactions will be the most favorable ones to have a CT reaction. Taking into account Table 1.3, it is necessary to highlight that phenylalanine is the most probable aromatic amino acid to form stacked interactions with guanine.

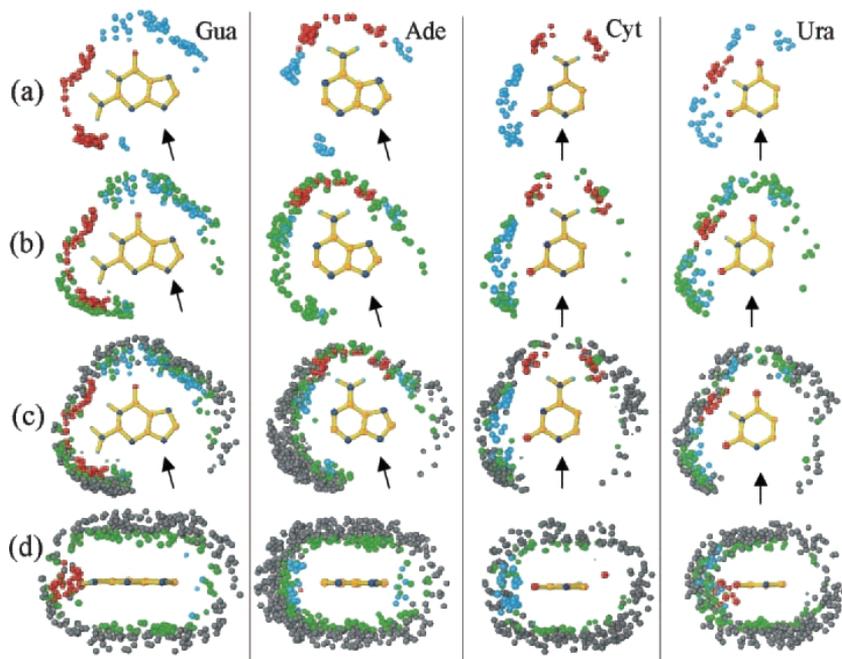


Figure 1.12: Composite nucleoside binding pockets produced by superposition of each base from 41 high resolution protein-RNA structures. Atoms from the protein making interactions to the base are shown as spheres: H-bond donors (blue); H-bond acceptors (red); VdW contacts (green); non-polar contacts (gray).⁶⁵

Table 1.3: Distribution of stacking interactions in 41 high resolution protein-RNA structures.

	Adenine	Uracil	Cytosine	Guanine
Phenylalanine	7	4	2	4
Histidine	6	3	0	0
Arginine	13	12	15	3
Tyrosine	5	3	2	1

1.3 Electron Transfer Theory

Rudolph A. Marcus developed a theory for electron transfer (ET) reactions for which he won the Nobel Prize in 1992.⁶⁶ This theory is well accepted because it can reproduce correctly experimental values of electron transfer rates obtained during the last decades.

Marcus theory is used to calculate electron transfer rates for systems consisting in two chemical species, a donor molecule and an acceptor one. This theory was formulated to address outer sphere electron transfer reactions. In these reactions the two chemical species only change their charge with an electron jumping or moving, without large structural changes. In these reactions, where coupling is weak, both donor and acceptor keep their identity during the reaction. Outer sphere ET reactions are those between nucleobases in DNA, as well as between NB and proteins.

Marcus theory was also extended to inner sphere electron transfer contributions, where the donor and acceptor species are connected by a chemical bridge. For inner sphere ET reactions, changes on the species bond distances are considered.

Outer sphere ET reactions present an activation energy. The rate for these reactions, controlled by activation energy, has an exponential form represented by equation 1.2.

$$k_{act} = A \cdot \exp\left(\frac{-\Delta G^\ddagger}{k_B T}\right) \quad (1.2)$$

In equation 1.2, ΔG^\ddagger is the Gibbs free energy of the formation of the transition state. The exponential term represents the probability of its formation and A contains the probability of crossing from precursor to successor complex.

The free activation energy can be calculated as shown in equation 1.3.

$$\Delta G^\ddagger = \frac{(\lambda_0 + \Delta G^0)^2}{4\lambda_0} \quad (1.3)$$

Where ΔG^0 is the Gibbs free energy change for the electron transfer reaction (see Figure 1.15) and λ_0 the reorganization energy. ΔG^\ddagger is

the intersection point between the parabolas that describe the outer-spheres reorganization energy of a system of two spheres in a solvent, Figure 1.13. Parabola i is the free energy surface of the initial state while parabola f is the free energy surface of the final state.

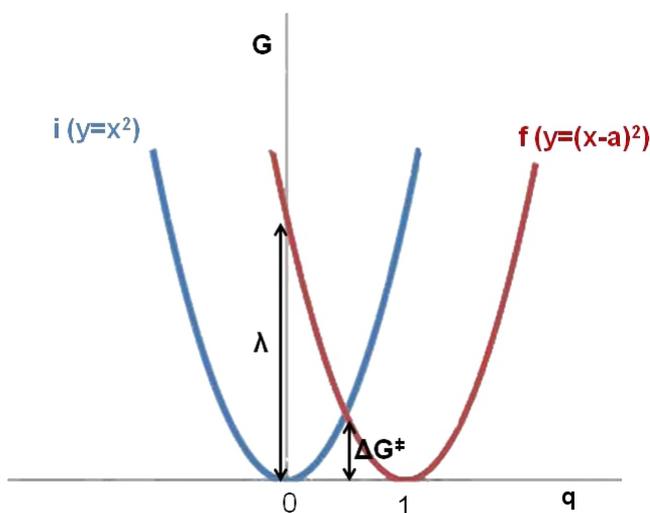


Figure 1.13: Parabolas of outer-sphere reorganization energy of the system two spheres in a solvent. The abscissa is the charge transfer reactions coordinate q while the ordinate is the Gibbs free energy.

The intersection of those parabolas represents an activation energy and not the energy of a transition state with a fixed configuration of all the nuclei of the system. Redox reactions do not have structurally determined transition states. Only energetic conditions are necessary. Thus, many polarization environments may meet these energetic conditions.

The dependence of the electron transfer rate on the reaction free energy is not direct. This is unusual because we generally think in terms of a linear free energy relationship between the rate logarithm of a reaction and the equilibrium constant. This leads to think that the rate should increase as we increase the driving free energy for the reaction ($-\Delta G^0$). This behavior only holds for a small region of outer-sphere redox reactions. Instead, equation 1.2 shows that the ET rate will increase with $-\Delta G^0$, until a maximum rate is observed for $\Delta G^\ddagger=0$,

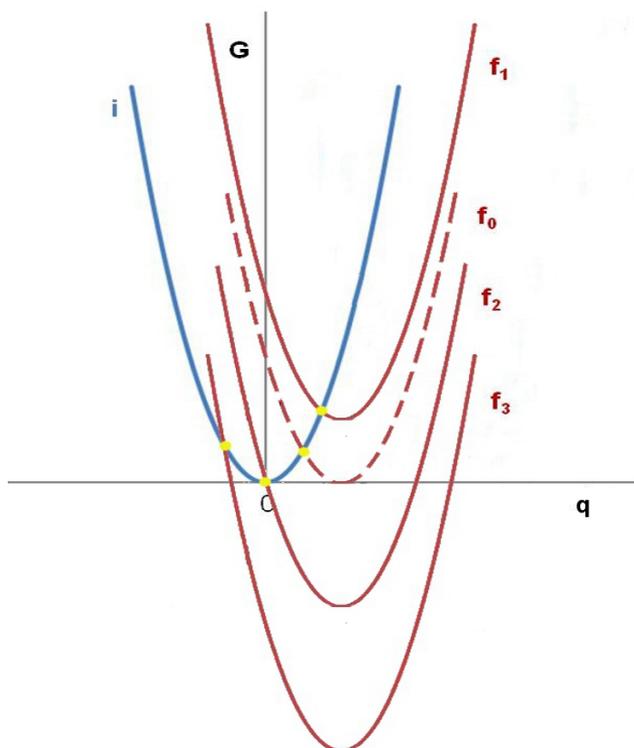


Figure 1.14: Marcus parabolas for different redox reactions: f_1 such with positive ΔG^0 , f_0 for the self-exchange reaction with $\Delta G^0=0$ (broken line), f_2 for moderately negative ΔG^0 (selected so, that $\Delta G^\ddagger=0$) and f_3 for strongly negative ΔG^0 . The free energy of activation ΔG^\ddagger (yellow point) decreases from f_1 via f_0 to f_2 (zero) and increases again for f_3 (Marcus inverted region).

and then the rate decreases. This decrease of the rate with increased ΔG^0 is called "Marcus inverted region". Figure 1.14 shows different examples of ΔG^\ddagger possible values (yellow points) depending on the ΔG^0 of the system. The inverted region was searched experimentally for 30 years until it was unequivocally verified in 1984.

The main equation of Marcus theory (Eq. 1.4) is used to calculate the electron transfer rate (k_{ct}).

$$k_{ct} = \frac{2\pi}{\hbar} |V_{da}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(\frac{-(\Delta G^0 + \lambda)^2}{4\lambda k_B T}\right) \quad (1.4)$$

Where k_{ct} stands for the electron transfer rate, $|V_{da}|$ represents the electronic coupling between the initial and final states, λ is the reorganization energy (both inner and outer-sphere), ΔG^0 is the Gibbs free energy change for the electron transfer reaction, and k_B is the Boltzmann constant.

Equation 1.4 is completely valid in the weak coupling limit which ensure a non-adiabatic regime, $V_{da} < kT \simeq 0.026$ eV. A high temperature is also required to employ this classical description of charge transfer. At low temperatures modules coupled to electron transfer must be treated as a quantum mechanical if the spacing between the vibrational levels is large compared to thermal energies. At low temperatures quantum mechanical tunneling through the barrier has a considerable influence in the charge transfer rate.⁶⁷⁻⁶⁹

For neighboring nucleobases in DNA electronic coupling is around 0.1 eV but the coupling to a second neighbor is several times weaker. Because of this, when the acceptor is separated from the donor by one or more base pairs we can consider the system within a non-adiabatic regime.⁷⁰ However, systems in an adiabatic regime, with strong couplings, will stay in the low potential energy parabola, see Figure 1.15.⁷¹

There are three parameters that determine the reaction rate in equation 1.4 which are represented in Figure 1.15.

V_{da} Donor-acceptor electronic coupling. This parameter determines the dependence of k_{ct} on the relative position between donor and acceptor sites. If there is a bridge between donor and acceptor sites, even if it is only a base pair, there is not a direct coupling, and then the reaction is mediated by superexchange. Actually those involved in the charge transfer process are the virtual states of the bridge.⁷²

ΔG^0 The driving force. It represents the total Gibbs free energy change of the charge transfer reaction. It also stands for the difference between redox potentials of the donor and acceptor sites. In a charge shifting case were $d^+ba \rightarrow dba^+$, ΔG^0 is independent of the distance (b =bridge) between d and a . The necessary energy for hole transfer between two molecules, can be estimated as the energy difference between their ionization

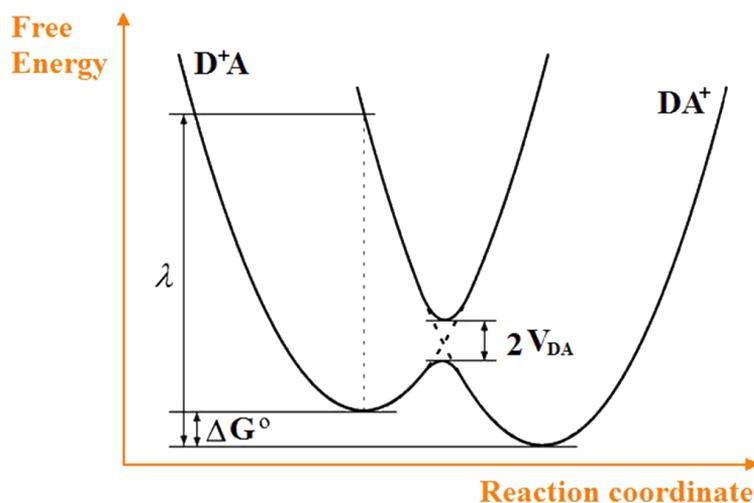


Figure 1.15: Potential energy surfaces of initial (left parabola) and final (right parabola) states in Hole Transfer.

potentials (ΔE). This happens because the holes are trapped in the minimum of oxidation potentials. As no important variations in the entropy of the systems studied in this thesis are expected, ΔG^0 can be approximated by ΔE .

λ Reorganization energy. According to the IUPAC, the reorganization energy is the energy required for all structural adjustments (in the reactants and the surrounding solvent molecules) which are needed in order that donor and acceptor sites assume the configuration required for the transfer of the electron. The reorganization energy is composed by two terms, the internal and solvent terms: $\lambda = \lambda_i + \lambda_s$.

1.3.1 Charge transfer parameters

Driving Force

Considering DNA-protein interactions, environmental effects on ΔG^0 should be similar to the ones observed for CT in DNA. The counterions of DNA environment strongly modulate the ionization potential of nucleobases, thus controlling the charge transfer rate. The redox potential, influenced by these counter-ions, fluctuates with a characteristic time of 0.3-0.4 ns as is shown in the work of Voityuk

et al.,¹⁶ (Figure 1.16). These fluctuations can even change the mechanism of charge transfer from G-Hopping to A-Hopping or gate the charge transfer. Notice that changes in the π stacking of nucleobases do not produce such changes in the driving force.^{73,74}

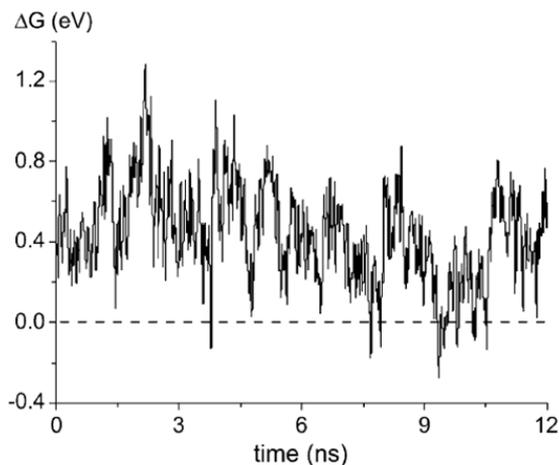


Figure 1.16: Fluctuations of the driving force (ΔG) for a hole transfer from G_3 to A_6 in the duplex 5'-TTG₃T₄T₅T₆T₇ ...-3', calculated along a MD trajectory of 12 ns.¹⁶

Electronic Coupling

This thesis is focused on the calculation of the electronic coupling for nucleobase-amino acid systems because electronic coupling strongly modify the rate of the charge transfer reaction. In chapter 2, an extensive explanation about electronic coupling calculations can be found.

For CT reactions in DNA, the environment of each nucleobase affects the rate of the reaction. The electronic coupling between nucleobases in the bridge has specificity and directional asymmetry. Purine bases are affected by electrostatic and exchange interactions with pyrimidine bases.⁷⁵ So as, by changing the composition of nucleobase bridge it is possible to control de rate of charge transfer. Fundamentally, it is possible to modify the virtual states of the bridge and the electronic

coupling with the adjacent bases.⁷⁶

Even though this work is based in the study of hole transfer, it is interesting to say that electronic coupling in excess electron transfer (*EET*) and hole transfer can be calculated similarly. In *EET*, anionic radicals of nucleobases are formed. Their treatment in quantum mechanics is more difficult than for neutral or cationic systems. The one electron approximation will fail to describe this negatively charged system, and the basis set used can greatly influence results. Instead of describing the excess charge state as the HOMO of the radical anion (one-electron approximation), it is described using the LUMOs of the neutral system calculated without diffuse functions. These orbitals can be used to estimate the *EET* coupling matrix elements. Generally coupling for excess electron transfer are much lower than couplings for hole transfer.⁷⁷

Biological systems, such as DNA and DNA-protein interactions, suffer structural fluctuations which affect the coupling between the molecules, Figure 1.17. In DNA case, in a time of 10 ps the matrix element squared can change 2-3 orders of magnitude depending on the length of the bridge. As electronic coupling is extremely sensitive to conformational changes, when CT reactions can be activated by switching between conformations a **gated mechanism** is observed. In such gated CT reactions the charge is not gradually transferred from donor to acceptor but sudden conformational changes of the system can enhance the reaction rate.^{20,78,79} This structural dependence can be solved by numerically integrating the charge transfer rate along the MD for a period around 1ns.^{80,81}

Reorganization energy

As stated in charge transfer theoretical basis section, reorganization energy consists of two terms, the internal and the solvent one: $\lambda = \lambda_i + \lambda_s$.

The internal term describes the intra-molecular geometric changes, both of the donor and the acceptor sites, which are caused by the charge transfer reaction. The internal reorganization term is usually estimated using quantum chemical methods.⁸²

For a hole transfer reaction like $d^+ + a \rightarrow d + a^+$, the internal reorga-

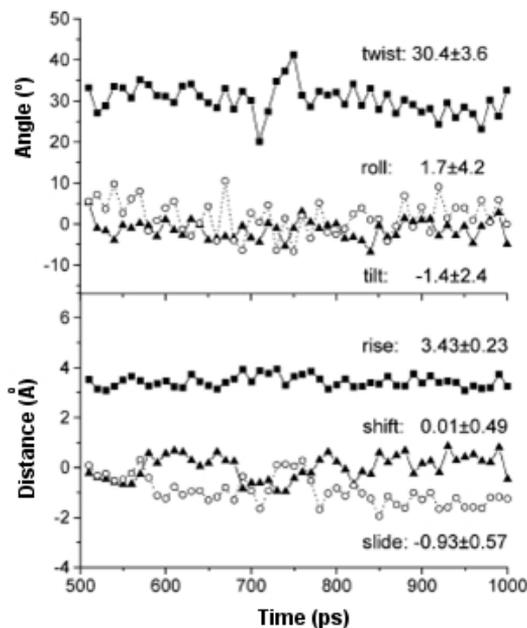


Figure 1.17: Fluctuations of conformational parameters of the base steps A_3A_4 in the duplex 5'-CCAACGTTGG-3' extracted from a MD trajectory of 1 ns.⁸⁰

nization energy is the sum of the donor and acceptor reorganization energy:

$$\lambda = \lambda_i(d) + \lambda_i(a).$$

Each of these terms can be expressed as:

$$\lambda_i(d) = E^+(d) - E^0(d)$$

$$\lambda_i(a) = E^0(a^+) - E^+(a^+)$$

For the donor site, $E^+(d)$ is the calculated optimized geometry for the oxidized state, and $E^0(d)$ is the calculated optimized geometry for the neutral state. In the acceptor case, $E^0(a^+)$ is calculated using the optimized geometry of neutral a and $E^+(a^+)$ uses the optimized geometry of the cation a^+ . Graphically λ_i can be estimated from Figure 1.15 but if the parabolas do not have the same curvature it has to be estimated as an average.

The charge transfer reaction produces slow changes in the polarization of the solvent. There are different approaches to determine the solvent term λ_s . The solvent reorganization energy strongly depends on calculation details. It is affected by the delocalization between nucleobases that decreases it. DNA structural fluctuations affect it (weakly). λ_s depends on the distance between donor and acceptor sites and increases the falloff parameter β that determines the conductivity of the bridge.^{44,82,83}

Several studies use different methods to estimate the solvent reorganization energy. LeBard et al.⁸⁴ report calculation of the solvent reorganization energy of charge transfer in DNA fragments based on the nonlocal molecular-based description of the solvent response. These authors apply the Q-model,⁸⁵ which is a quadratic coupling model where the Hamiltonian of a two-state solute is linearly coupled to a harmonic solvent mode with different force constants in the initial and final states. Q-model predicts a much lower sensitivity of the overall rate constant to changes in the solvent reorganization energy than traditional theories.

1.3.2 Excess Charge Delocalization

The excess charge delocalization between donor and acceptor has a very significant role in charge transfer processes. It depends on different parameters related to each other, electronic coupling between donor and acceptor sites, internal reorganization energy and interactions with polar environment. Different theoretical mechanism can be used to study the propagation of localized and delocalized states. However, it still has points to solve and different studies have shown controversial results.^{86,87} For example, charge transfer through DNA π stacking depends on whether the charge is delocalized or not. With a better delocalization, more extended wave function, a better conductance is obtained. For hole transfer processes, several studies use mechanisms that imply the hole delocalization between several nucleobases of DNA stacking.^{88,89}

So far there has been no experimental evidence of charge delocalization in DNA. This may be caused by the polar environment.^{44,90} However, there are several mechanisms implying charge delocalization:

- Polaron model: The insertion of a charge in DNA produces

changes in its structure and in the environment. This happens because the timescale of charge hopping is similar to the timescale of the DNA and environment dynamics. These changes induce the formation of a polaron which is an energy minimum of the system. The radical cation is self-trapped by this minimum and needs of thermal energy to overcome the energy barrier.^{6,91,92}

- Ion-gated model: In this mechanism, the delocalized charge is controlled by DNA's polar environment (water and sodium ions). This model has been combined with the polaron model in several studies.^{92,93}
- Domain model: A domain is a transiently extended π -orbital defined dynamically by DNA sequence over which charge can delocalize. A domain may include the hole acceptor, or may dynamically couple with a domain that includes the acceptor. A domain, or collection of domains modulating electronic communication, constitutes a CT-active conformation.⁸⁹

Chapter 2

Methodology

The Methodology chapter is divided mainly in three parts:

The first section is an overview of different methods to calculate electronic coupling. The calculation of this parameter is one of the main objective of this thesis because of its influence on charge transfer rates.

Then following three sections briefly describe the *ab initio* methodologies that have been used in this thesis. *Ab initio* methods do not use any experimental data in their calculations. The structures and energies of the molecules are obtained by solving the non-relativistic time-independent Schrödinger equation.⁹⁴ However, the Schrödinger equation can only be explicitly solved for the hydrogen atom. Several approximations have been developed to face larger systems. First, density functional methods will be overview. Second multi-configurational methods are described. Then the inclusion of the dynamic correlation (Møller-Plesset perturbation theory) in the multistate calculations is explained. Finally, several methods to calculate electronic couplings are explained in detail. Much more information of the basis of *ab initio* methodologies can be found in the books by Szabo,⁹⁵ Helgaker,⁹⁶ Yarkony,⁹⁷ and Jensen.⁹⁸

The last section of this chapter is a short explanation about the structural nomenclature that has been employed.

2.1 Calculation of electronic coupling

This section presents several methods to calculate electronic coupling values (V_{da} or H_{da}). As it was said before, V_{da} is one of the key parameters that determines the rate of the charge transfer reaction, so this thesis is mainly focused on its calculation. However, driving force and reorganization energy should be also taken into account.

2.1.1 Direct Method

The direct method has been largely used in this thesis, both in the study of NB-*aaa* interactions as well as in the analysis of conformational parameters in the coupling between guanine-tryptophan and adenine-tryptophan systems (see. chapters 4, 5, and 6). In many cases, this computational scheme provides reliable estimations of electronic coupling and it is widely employed to study HT in DNA.⁹⁹ Our tests have shown that it is more robust than other 2-state models such as generalized Mulliken-Hush (GMH) or Fragment Charge Difference method (FCM) (see sections 2.1.2 and 2.1.3).

Two diabatic states of donor (d) and acceptor (a) sites are considered, ψ_d and ψ_a . The wave functions to calculate the effective Hamiltonian matrix elements (\mathbb{H}') and the overlap integrals (S_{da}) are represented by the ground-state wave functions of the isolated d and a . Thus, this method uses a standard diabatic representation in which the initial and final electronic states are by construction taken as the valence bond structures corresponding, respectively, to the reactants (ψ_d) and products (ψ_a) of the reaction.

The electron transfer integral (H_{da}) is defined as:

$$H_{da} \equiv \langle \psi_d | H_{el} | \psi_a \rangle \quad (2.1)$$

where H_{el} is the electronic Hamiltonian.

The effective Hamiltonian matrix, \mathbb{H}' , for the case of nonorthogonal diabatic states ($S_{da} \neq 0$) is:

$$\mathbb{H}' = \begin{pmatrix} 1 \\ 1 - S_{da}^2 \end{pmatrix} \begin{pmatrix} H_{dd} - S_{da}H_{ad} & H_{da} - S_{da}H_{aa} \\ H_{ad} - S_{ad}H_{dd} & H_{aa} - S_{ad}H_{da} \end{pmatrix} \quad (2.2)$$

In the limit of small S_{da} , symmetric orthogonalization yields (via Taylor expansion) to equation 2.3.

$$H'_{da} = H'_{ad} \simeq H_{da} - S_{da}(H_{dd} + H_{aa})/2 \quad (2.3)$$

Where H'_{da} is the electronic coupling between the donor and acceptor states.⁹⁹

2.1.2 2-state Generalized Mulliken-Hush

For a system of two well-separated molecules (donor and acceptor), the 2-state model can be applied. However, Voityuk's previous studies⁹⁰ have shown that, if there is a bridge between the donor and acceptor sites (like an adenine nucleobase in a CT reaction between guanines in DNA), it is preferable to use the multistate generalized Mulliken-Hush (GMH) method. This method was developed by Cave and Newton.^{72,90,100,101} This model employs a transformation of adiabatic states into diabatic states, that diagonalizes the adiabatic dipole moment matrix. The electronic coupling, off-diagonal matrix elements of the diabatic Hamiltonian, can be calculated as:

$$V_{da} = \sum \mathbb{T}_{id} \mathbb{E}_i \mathbb{T}_{ia} \quad (2.4)$$

T→Unitary transformation.

E→Diagonal matrix of the adiabatic energies.

Applying the 2-state model, electronic coupling can be expressed via the vertical excitation energy, also called adiabatic splitting ($E_2 - E_1$), the difference of the adiabatic dipole moments ($\mu_1 - \mu_2$), and the transition dipole moment (μ_{12}).

$$V_{da} = \frac{(E_2 - E_1)|\mu_{12}|}{\sqrt{(\Delta\mu_2 - \Delta\mu_1)^2 + 4\mu_{12}^2}} \quad (2.5)$$

The minimum splitting is achieved when donor and acceptor are "in resonance" which implies $\mu_1 = \mu_2$ in equation 2.5 and $\Delta\mu = 0$ in equation 2.7.

$$V_{da} = \frac{1}{2}(E_2 - E_1)\sqrt{1 - \Delta\mu^2} \quad (2.6)$$

$$V_{da} = \frac{1}{2}(E_2 - E_1) \quad (2.7)$$

As shown in equation 2.7, when the splitting ($E_2 - E_1$) is small, V_{da} can be estimated as a half of this splitting.

The GMH method allows calculation of H_{da} in various systems, independent of symmetry and geometric constraints. The GMH method can deal with multistate situations where more than two adiabatic states enter into the description of the diabatic states of interest. The details for a 3-state calculation are explained in section 2.1.4.

Electronic coupling can be calculated using the 2-state model if there is only donor and one acceptor site without a bridge between them. Systems with a bridge have been calculated using Hartree-Fock level.^{70,72} However, the 2-state model fails to estimate H_{da} for systems where the bridge is energetically close to donor and acceptor sites. This happens even for systems with remarkably small bridges. To improve results, it is better to use the GMH method and check if the results obtained with the 2-state model are consistent.^{90,102}

2.1.3 Fragment Charge Method

Another method to estimate the electronic coupling is the fragment charge method (FCM).¹⁰³ V_{da} can be estimated using FCM with an analogous of equation 2.5:

$$V_{da} = \frac{(E_2 - E_1)|q_{12}|}{\sqrt{(\Delta q_2 - \Delta q_1)^2 + 4q_{12}^2}} \quad (2.8)$$

In equation 2.8, Δq_1 and Δq_2 are d - a charges difference in the adiabatic states Ψ_1 and Ψ_2 , respectively, of the 2-state model and Δq_{12} is the corresponding off-diagonal term. Ψ_1 and Ψ_2 are the two orthonormalized adiabatic states that describe the system and correspond to the normalized diabatic wave functions of the donor and acceptor sites, φ_d and φ_a .

2.1.4 3-state GMH

Usually on HT studies involving the indole ring, it is assumed that the 2-state model is a suitable approximation to calculate electronic coupling, but to our knowledge such an assumption has not been tested. Compared to the systems usually considered in HT reactions inside DNA (π -stacked nucleobases) for which the 2-state model can be employed, the (G-Ind) \bullet^+ system (the most studied interaction in

this thesis) is particularly challenging. This is due to the proximity in energy of ground state and first excited state of Indole radical cation. In such a case, the 2-state model to calculate the electronic coupling of the system becomes questionable, and the extension to a 3-state model may become necessary.^{90,101,104,105}

When is the 3-state GMH required?

A simple diagnostic method proposed by Cave et al.¹⁰⁶ can be used to predict when a third state should be considered to calculate the electronic coupling element for a given pair of diabatic states, within the context of the Generalized Mulliken-Hush approach. The diagnostic is formulated on the basis of Löwdin partitioning theory.

For the 3-state results discussion, we will consider a model having a pair of diabatic states with the cationic charge or hole localized on the acceptor molecule (e.g. the ground state (GS) of the acceptor and its locally excited (LE) state) and a single charge-transfer state (CT), having the hole localized on the donor molecule. Figure 2.1 represents the electronic model corresponding to $(\text{G-Ind})^{\bullet+}$ system, where Ind presents two different electronic states (GS and LE) while G only Ground state.

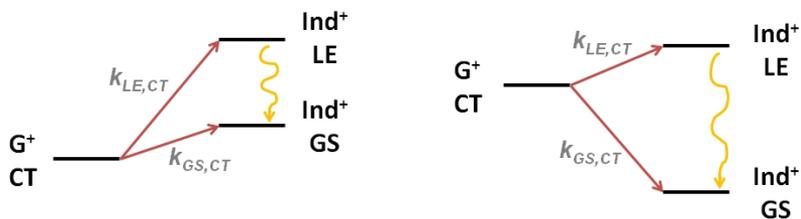


Figure 2.1: For $(\text{G-Ind})^{\bullet+}$ systems 3 states can be involved in the charge transfer reaction. One from Guanine (donor) and two from Indole (hole acceptor).

Thus, applying Cave diagnostic, the need of a 3-state method can be derived from the transition dipole moments of the adiabatic states:

$$\left| \frac{\frac{\mu_{GS,LE}\mu_{LE,CT}}{\mu_{CT} - \mu_{LE}}}{\mu_{GS,CT}} \right| = |\lambda_D| \ll 1 \quad (2.9)$$

When λ_D value is much less than unity, the multistate effect will be negligible. Thus, a 2-state model can be applied.

This diagnostic method has been employed during this thesis (see chapter 7) and has been shown to be very useful and accurate. Notice that the necessity to employ the 3-state method strongly depends on the conformation between the studied fragments (G and Ind in our case). It is worth noting that 2-state or 3-state treatment will depend on conformational changes.

3-state GMH calculations of electronic coupling

As it was introduced before, HT reactions between G and Ind can be treated explicitly as a 3-state case using the multistate GMH.

Multi-referential methods are necessary to obtain the wave function of the adiabatic electronic states in (G-Ind) $^{\bullet+}$ complex. For example, the use of the MS-CASPT2 method (see section 2.4.2) will lead to a good description of the adiabatic states.

To calculate electronic couplings, first the projection of the dipole moments along the charge transfer vector must be done. The charge transfer vector is defined as the vector connecting the centers of mass of the two molecules. Then, the adiabatic dipole moment matrix (\mathbb{M}_{ad}) can be diagonalized and the \mathbb{M}_{diab} matrix is obtained:

$$\mathbb{U}^T \mathbb{M}_{ad} \mathbb{U} = \begin{pmatrix} \mu_{GS}^{diab} & 0 & 0 \\ 0 & \mu_{CT}^{diab} & 0 \\ 0 & 0 & \mu_{LE}^{diab} \end{pmatrix} \quad (2.10)$$

Notice that the dipole moments that belong to the two states of Ind approximately should have the same value (μ_{GS} and μ_{LE}). The third one (μ_{CT}), with a different value, belongs to the guanine state.

The eigenvector matrices \mathbb{U} and \mathbb{U}^T are obtained during the diagonalization of \mathbb{M}_{ad} . These eigenvector matrices can transform the adiabatic energy matrix \mathbb{E} into the diabatic Hamiltonian matrix \mathbb{H} .

$$\mathbf{U}^T \mathbb{E} \mathbf{U} = \begin{pmatrix} H_{GS} & V_{GS,CT} & V_{GS,LE} \\ V_{CT,GS} & H_{CT} & V_{CT,LE} \\ V_{LE,GS} & V_{LE,CT} & H_{LE} \end{pmatrix} \quad (2.11)$$

Where the off-diagonal terms represent the electronic coupling (V) between the states.

Up this point, the explained methodology is common for the 2-state calculation. Considering a third state in the calculation differs hereafter.

As the electronic coupling concerns, it has physical meaning only in intermolecular interactions; the coupling between the two states from Indole must be zeroed. In order to achieve this, a block diagonalization matrix (\mathbb{B}) is used. The diabatic Hamiltonian matrix (\mathbb{H}) is then transformed into a pseudo-diabatic Hamiltonian matrix (\mathbb{H}').

$$\mathbb{B}^T \mathbb{H} \mathbb{B} = \mathbb{H}' = \begin{pmatrix} H_{GS} & V_{GS,CT} & 0 \\ V_{CT,GS} & H_{CT} & V_{CT,LE} \\ 0 & V_{LE,CT} & H_{LE} \end{pmatrix} \quad (2.12)$$

In this pseudo-diabatic Hamiltonian matrix, the off diagonal terms correspond to the electronic coupling between guanine and the two states of Ind ($V_{GS,CT}$, $V_{LE,CT}$). As can be seen the $V_{GS,LE}$ coupling between the Ind states has been stated as 0.

For 3-state cases the k_{HT} will be the sum over two possible hole transfer reactions of guanine (CT) with the two states of indole (GS, LE). The hole transfer reaction rate can be written as:

$$k_{HT} = k_{GS,CT} + k_{LE,CT}$$

If a strong coupling between CT and LE states exist, and the driving force between them is small, high $k_{LE,CT}$ values are obtained. Hence, the k_{HT} can be drastically modified.

2.1.5 Koopmans' Theorem Approximation

Koopmans' Theorem Approximation (KTA) states that in closed-shell Hartree-Fock theory, the first ionization energy of a molecular system is equal to the negative of the orbital energy of the highest occupied molecular orbital (HOMO).

The radical cation states D^+A and DA^+ represent the initial and final states of the CT reaction between donor and acceptor molecules. The adiabatic splitting of the electronic states is $\Delta=E_2-E_1$. Using KTA, Δ can be estimated as the difference of the one-electron energies of the two highest occupied molecular orbitals HOMO and HOMO-1 calculated for the closed-shell neutral system DA. Within this approximation, distribution of the excess charge in the ground state of the radical cation can be estimated via the corresponding Mulliken populations of the HOMO of the neutral system. Then, the charge on a fragment F can be estimated as:

$$q_1(F) = \sum_{i \in F} C_{i,HOMO} \sum_{j=1}^N C_{j,HOMO} S_{ij} \quad (2.13)$$

Here, S_{ij} is the overlap of atomic orbitals (AOs) i and j ; i runs over atomic orbitals associated with the fragment F while j runs over all AOs. The fragment charges in the first excited state are calculated similarly using the coefficients HOMO-1.

Koopmans' approximation can be used with the GMH method (see section 2.1.2). Within KTA the difference of the adiabatic dipole moments ($\mu_1-\mu_2$) and the transition moment μ_{12} can be defined as:

$$\mu_1 - \mu_2 = \sum_{ij=1}^M (C_{i,HOMO}C_{j,HOMO} - C_{i,HOMO-1}C_{j,HOMO-1})d_{ij} \quad (2.14)$$

$$\mu_{12} = \sum_{ij=1}^M C_{i,HOMO}C_{j,HOMO-1}d_{ij} \quad (2.15)$$

Here, d_{ij} are the matrix elements of the dipole operator defined for AOs i and j .

2.2 Density Functional Theory

The Density Functional Theory (DFT) is based on the Hohenberg and Kohn¹⁰⁷ theorems which demonstrated that the ground-state electronic energy is determined completely by the electron density ρ . However, the functional which connects ρ and energy is not known. Before Hohenberg-Kohn demonstration, electronic density also plays a central role in Thomas-Fermi^{108,109} method. In this method, electrons are represented in a similar way to a classical liquid. In contrast, Hohenberg-Kohn formulated DFT as an exact theory of many-body systems. DFT is based in the two theorems:

- Theorem 1: The ground state electron density is sufficient to construct the full Hamilton operator and hence to calculate - in principle - any ground state property of the system without the knowledge of the many electron wave function. Alternatively formulated, this means that any ground state property can be expressed in terms of the ground state electron density ρ_0 . The ground state energy of a system E_0 , is a unique functional of the ground state density. $E_0 = E[\rho_0(r)]$
- Theorem 2: The functional for the ground state energy is minimized by the ground state electron density. The energy of the system is minimum when the exact density of the system ρ_0 is considered.

$$E[\rho] \geq E[\rho_0] \quad (2.16)$$

Applying this, E_0 can be found by minimizing $E[\rho]$ with a variational method:

$$E[\rho] = F_{HK} + \int v_{ext}\rho(r)dr \quad (2.17)$$

where v_{ext} stands for the external potential which represents the electron-nuclei attraction:

$$v_{ext} = \sum_{a=1}^M \frac{-Z_a}{|-R_a - r|} \quad (2.18)$$

F_{HK} is a universal function of electron density, which depends on kinetic energy ($T[\rho]$), the classical Coulomb energy ($J[\rho]$) and on the non-classical electron-electron interaction energy ($E_{NC}[\rho]$):

$$F_{HK}[\rho] = T[\rho] + J[\rho] + E_{NC}[\rho] \quad (2.19)$$

2.2.1 Kohn-Sham equations

The Kohn-Sham equations allow to find the ground state density. The main advantage of this method is that it handles kinetic energy easily by employing a fictitious non-interacting electron system to mimic the true many-electron system. Hence, the ground state wave function can be written in terms of basic one-electron orbitals. Even so, the full wave function (Ψ_S) have to satisfy exchange-antisymmetry. This can be done by placing the one-electron wave function ψ_i , in an Slater determinant:

$$\Psi_S = \frac{1}{\sqrt{N!}} \det\{\psi_i(x_1)\psi_j(x_2)\dots\psi_k(x_N)\} \quad (2.20)$$

In his work of 1965, Kohn and Sham¹¹⁰ divided $E[\rho]$ in four parts:

$$E[\rho] = T_s[\rho] + \int \rho(r)v_{ext}(r)dr + J[\rho] + E_{XC}[\rho] \quad (2.21)$$

- $T_s[\rho]$ is the kinetic energy of non-interacting electrons.
- v_{ext} is the external potential, see eq. 2.18.
- $J[\rho]$ is the classical Coulomb energy.
- $E_{XC}[\rho]$ is the exchange-correlation energy which includes the non-classical electron-electron interaction energy and the kinetic energy corresponding to the fully interacting system $T[\rho]$.

Electronic density can be represented as a set of occupied one-electron orthonormal orbitals $\{\psi_i\}$. A functional of the orbitals can be defined as:

$$\Omega[\psi_i] = E[\rho] - \sum_i \sum_j \varepsilon_{ij} \int \psi_i^*(r)\psi_j dr \quad (2.22)$$

ε_{ij} are Lagrange multipliers to ensure that the orbitals are orthonormal.

The minimization of $\Omega[\psi_i]$ respect $\psi_i^*(r)$ gives the *Kohn-Sham equations*:

$$\left[-\frac{1}{2}\nabla_i^2 + v_{eff}(r) \right] \psi_i(r) = \varepsilon_i \psi_i(r) \quad (2.23)$$

$v_{eff}(r)$ is the Kohn-Sham potential defined as:

$$v_{eff}(r) = v_{ext}(r) + v_H + v_{xc}(r) \quad (2.24)$$

- v_H is the Hartree potential.

$$v_H = \int \frac{\rho(r')}{|r - r'|} dx' \quad (2.25)$$

- And $v_{xc}(r)$ is the exchange-correlation potential.

$$v_{xc}(r) = \frac{\delta E_{xc}}{\delta \rho(r)} \quad (2.26)$$

Much effort have been done to find accurate approximations of v_{xc} because the accuracy of DFT calculations depends considerably on it.

2.2.2 DFT Methods

There are three different types of density functional methods.

- Local Density Approximation (LDA) methods assume that the density of the molecule is uniform throughout the molecule. Initially Kohn Sham implementations used this approximation but nowadays is hardly used.
- Gradient-Corrected (GC) methods look to account for the non-uniformity of the electron density.
- Hybrid methods attempt to incorporate some useful features from ab initio methods (specifically Hartree-Fock methods) with some of the improvements of DFT mathematics. Hybrid methods, such as B3LYP, tend to be the most commonly used methods for computational chemistry practitioners.

Local Density Approximation methods

The Local Density Approximation (LDA) introduced by Kohn and Sham¹¹⁰ is valid for slowly varying density. It assumes that density can be treated locally as a uniform electron gas. Using LDA the exchange correlation energy can be expressed as:

$$E_{xc}^{LDA} = \int \rho(r) \varepsilon_{xc}^{hom}(r) dr \quad (2.27)$$

In eq. 2.27, ε_{xc}^{hom} stands for the exchange-correlation energy per particle of an uniform gas of density ρ . The exchange correlation potential can be described as:

$$V_{xc}^{LDA}[\rho] = \frac{\delta E_{xc}^{LDA}}{\delta \rho} = \varepsilon_{xc}^{hom}[\rho] + \rho \frac{\delta \varepsilon_{xc}^{hom}[\rho]}{\delta \rho} \quad (2.28)$$

ε_{xc}^{hom} can be split into an exchange (ε_x^{hom}) and correlation (ε_c^{hom}) potentials. The exchange potential for a uniform electron gas is given by the Dirac formula:

$$\varepsilon_x^{hom}[\rho] = -C_x \rho^{1/3} \quad (2.29)$$

where,

$$C_x = \frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} \quad (2.30)$$

LDA has been replaced by the Local Spin Density Approximation (LSDA), which is given as the sum of the individual densities raised to the 4/3 power. The exchange energy is given by:

$$E_x^{LSDA} = -2^{1/3} C_x \int [\rho_\alpha^{4/3} + \rho_\beta^{4/3}](r) dr \quad (2.31)$$

And the exchange potential is:

$$\varepsilon_x^{LSDA} = -2^{1/3} C_x [\rho_\alpha^{1/3} + \rho_\beta^{1/3}] \quad (2.32)$$

As stated before, nowadays LDA and LSDA methods are hardly used. The LSDA approximation in general underestimates the exchange energy by $\sim 10\%$, it could create errors which are larger than the whole correlation energy. Electron correlation is furthermore overestimated, often by a factor close to 2, and bond strengths are as a consequence also overestimated. However, despite the simplicity of the fundamental assumptions, LSDA methods are often found to provide similar results than Hartree-Fock methods.

Gradient-Corrected methods

LDA approximates the real density to a local constant density. However, in the molecules the density changes quickly, so LDA produces considerable errors. To improve the density calculation a non-uniform electron gas has to be considered. One way to achieve this is to make the exchange and correlation energies dependent not only on the electron density, but also on derivatives of the density. These methods are known as Gradient Corrected or Generalized Gradient Approximation (GGA) methods.

Different kinds of functions are used in GGA methods to modify the behavior of large gradients without changing the desired properties of the systems. Generally they can be described as:

$$E_{xc}^{GGA} = \int \rho(r) \varepsilon_{xc}^{hom} F_{xc}[\rho, \nabla \rho] dr \quad (2.33)$$

where,

- $\nabla \rho$ is the gradient of the density.

- F_{xc} is dimensionless.
- ε_{xc}^{hom} is the exchange-correlation energy of the uniform electron gas.

Some characteristic GGA functionals are: *CAM* developed by Handy and coworkers,¹¹¹ *B88* developed by Becke,¹¹² *PW91* by Perdew, *PBE* by Perdew, Bruke and Ernzerhof,¹¹³ and the popular *LYP* functional.¹¹⁴

Hybrid methods

The exchange-correlation energy and the corresponding potential can be connected by linking the non-interacting reference and the actual system. The resulting equation is called the Adiabatic Connection Formula (ACF)¹¹⁵ and involves a integration over λ parameter, which switches on the Coulomb electron-electron repulsion.

$$E_{xc} = \int_0^1 \langle \Psi_\lambda | V_{xc}(\lambda) | \Psi_\lambda \rangle d\lambda \quad (2.34)$$

or simplifying the nomenclature:

$$E_{xc} = \int_0^1 E_{xc}^\lambda d\lambda \quad (2.35)$$

A $\lambda=0$ stands for the non-interacting Kohn-Sham reference system. $\lambda=1$ will be the entire real interacting system. However, it is not practical to perform this integral thus equation 2.35 is approximated. The simplest way to do the approximation is by a linear interpolation (eq. 2.36):

$$E_{xc} = \frac{1}{2}(E_{xc}^0 + E_{xc}^1) \quad (2.36)$$

Another possibility to approximate 2.35 is by employing a 3-parameter equation. These parameters are semi-empirical and have been found by fitting the heats of formation for a standard set of molecules. One example of this method is the 3-parameter equation used by the *B3LYP* functional. *B3LYP* is one of most popular functionals and is also used in some calculations of this thesis.

$$E_{xc}^{B3LYP} = (1 - a)E_x^{LSD} + aE_x^{HF} + bE_x^{B88} + cE_c^{LYP} + (1 - c)E_c^{LSD} \quad (2.37)$$

Where, $a=0.20$, $b=0.72$, and $c=0.81$.

- E_x^{LSD} is the local spin density approximation to the exchange energy.
- E_x^{HF} is the Hartree-Fock exchange energy.

- E_x^{B88} is the Beck (B88) exchange functional.
- E_c^{LYP} is the LYP correlation functional.
- E_c^{LSD} is the local spin density approximation to the correlation energy.

2.3 Multi-Configurational Methods

Some systems studied in this thesis present locally excited states close in energy to their ground states. Then the general 2-state methods GMH and FCM to estimate electronic coupling (see section 2.1.4) can not be applied because the resulting states are mixed. In this way, the system is dominated by more than one electronic configuration. To obtain more accurate results, a multi-determinant method must be employed such as Multi-Configurational Self Consistent Field (MCSCF) methods. Moreover, in order to get a better description of the HT process the 2-state model must be replaced sometimes by a 3-state model.

The CASSCF and CASPT2 levels of theory should be more adequate for the calculation of the cationic systems studied in this thesis than use of time-dependent density functional (TD-DFT) methods. TD-DFT methods have limitations with the calculations of charge-transfer excitations.¹¹⁶

The basis of the MCSCF method is the Configuration Interaction (CI) method. It is founded on the variational principle, analogous to the HF method. The trial wave function is written as a lineal combination of Slater determinants, with the expansion coefficients determined by requiring that the energy should be a minimum or at least stationary.

A generic multi-determinant trial wave function can be written as:

$$\Psi_{CI} = a_0\Phi_{SCF} + \sum_S a_S\Phi_S + \sum_D a_D\Phi_D \dots = + \sum_{i=0} a_i\Phi_i \quad (2.38)$$

where, subscripts S, D, T ... indicate determinants which are singly, doubly, triply ... excited in relation to the HF configuration. Consequently the ground state of the CI method corresponds to the HF method one, which can be seen as a particular case of CI where $a_i = 0$ and $a_0 = 1$. In most cases the HF method gives a reasonable description of the wave function. The final CI wave function is composed in a 90% by the HF one. The Molecular Orbitals (MOs) used for building the excited Slater determinants are taken from a HF calculation and held fixed.

CI can give highly accurate wave functions for small systems. The number of Slater determinants needed for a system of n orbitals and $2k$ electrons is:

$$N_{det} = \binom{n}{k}^2 \quad (2.39)$$

Consequently the use of CI it is in most cases impossible. Thus, only truncated CI methods are of practical application. For example, CISD method which only takes into account configuration states made of single and double excitations.

However, truncated CI methods present two main problems:

- *No size-extensive*: For a size-extensive method, the energy calculated scales linearly with the number of particles.¹¹⁷ This allows to make comparisons between calculations with different number of electrons. For example, calculations with different number of active electrons. Truncated CI methods are not size-extensive, this also implies that the error of the energy estimation grows with the number of electrons used in the calculation.
- *Orbital generation*: The orbitals of the reference configuration state are generated self-consistently in the field of a single electronic configuration. Hence, their relevance can not be assured for a multi-configurational system.

The MCSCF method solves this second problem because not only the coefficients in front of the determinants are optimized by the variational principle, but also the MOs used for constructing the determinants are optimal. The MCSCF optimization is iterative. Since the number of MCSCF iterations required for achieving convergence tends to increase with the number of configurations included, the size of MCSCF wave function that can be treated is somewhat smaller than for CI methods.

The main difficulty with MCSCF methods is selecting the necessary configurations to be included for the property of interest. One of the most popular approaches of the MCSCF family is the Complete Active Space Self-consistent Field (CASSCF).

2.3.1 CASSCF

The Complete Active Space Self-consistent Field (CASSCF) method is also called Full Optimized Space (FORS). In this approach, the orbital space is divided in *active* and *inactive* orbitals. The active MOs will typically

be some of the highest occupied and some of the lowest unoccupied MOs. Thus, inactive orbitals corresponds to core orbitals (doubly occupied) and virtual orbitals (unoccupied), see Figure 2.2. The active space is composed by the active orbitals, these must be relevant to describe the studied process. For example, in our case, where the electronic coupling between the molecules determines the HT capability of the system, π orbitals must be in the active space. Within the active MOs, a full CI is performed and all the proper symmetry adapted configurations are included in the MCSCF optimization.

In a full CI calculation, the wave function is a lineal combination of Configuration State Functions, CSFs. CSFs are composed by the lineal combination of one or more Slater determinants which are eigenfunctions of spin operators. All Slater determinants (or CSFs) of the proper symmetry are included in the variational procedure (i.e. all Slater determinants obtained by exciting all possible electrons to all possible virtual orbitals, orbitals which are unoccupied in the electronic ground state configuration).

The suitable orbitals to be include in the active space must be selected manually. Hence CASSCF level of theory cannot be fully automatized. Usually, the notation used to reference the active space used is:

[n, m]-CASSCF

where n is the number of electrons which are distributed in all possible ways in m orbitals.

As can be seen in Figure 2.2, the CASSCF method only accounts for the electronic correlation of the orbitals within the active space. Hence, excitations only are allowed inside this active space. The active space must be selected carefully. As a full CI expansion, CASSCF becomes extremely large even for quite small active space. Our calculations were performed using an active space of (11,12). A larger active space of (14,15) was also tested but their computational cost was unaffordable to perform the HT study in our systems (up to 32 atoms). Nowadays CAS calculations are usually limited to a maximum active space size of (14,14).

The Restricted Active Space Self-consistent Field (RASSCF) method is a variation of CASSCF. As can be seen in Figure 2.2 the active MOs are divided in three sections: RAS1, RAS2 and RAS3. Each division have restrictions on the occupation number (excitations) allowed. Usually a HF reference determinant is employed to select the orbitals.

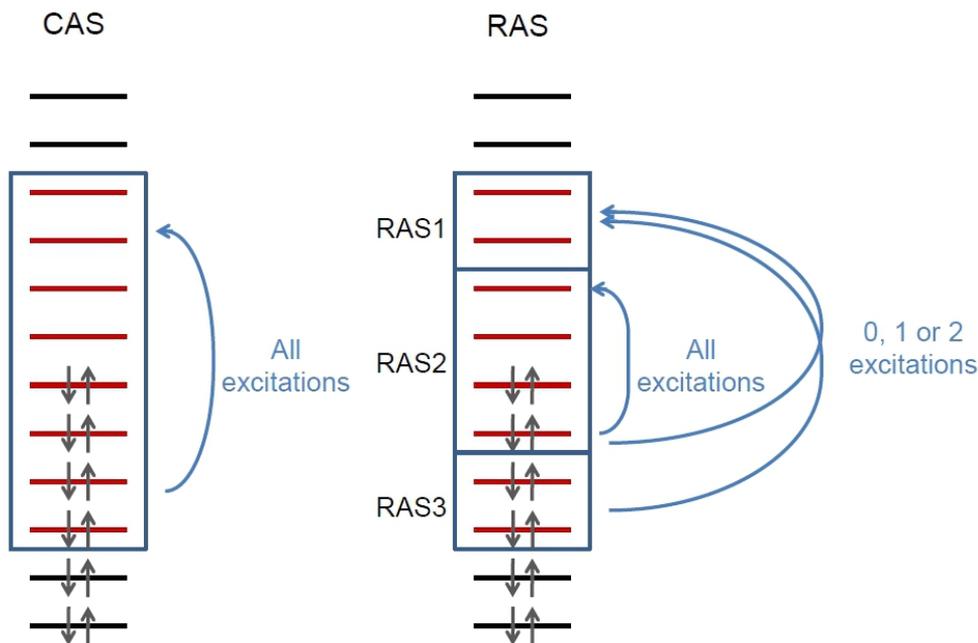


Figure 2.2: CAS and RAS MOs scheme.

There are few rules to select an appropriate active space:

1. Usually there will be the same number of occupied and virtual orbitals. This leads naturally to $[n, m]$ -CASSCF wave functions where n and m are identical or nearly so.
2. The orbital energies from a RHF calculation can be used to select the relevant orbitals. HOMO and LUMO are usually the most important orbitals to include in the active space. There are two situations when the use of the RHF energies could be problematic:
 - a) When extended basis sets are used many virtual orbitals can have low energies and the exact order is unknown.
 - b) If the real wave function has significant multi-configurational character, then the RHF wave function may be qualitatively wrong, and selecting the active orbitals based on a qualitatively wrong function may lead to erroneous results. The important orbitals to describe a multi-determinant nature are only really known after the multi-configurational calculation.

3. The type of reaction studied has to be considered in order to select the most suitable active space. Although, from the energetic point of view n-type orbitals should be included in the calculations performed in this thesis, charge transfer reactions are mainly described by π -orbitals. Thus, n-type orbitals have been replaced by π ones.
4. Natural orbitals can help to select the correct active space. They are those which diagonalize the density matrix, and the eigenvalues are the occupation numbers. Orbitals with occupation numbers significantly different from 0 or 2 (for a closed-shell system) are usually the most relevant to be included in the active space. To select these orbitals efficiently a MP2 calculation can be used.

To see how the electronic wave function of CASSCF is constructed or further details the books of Helgaker⁹⁶ and Yarkony⁹⁷ can be consulted.

2.4 Including dynamic electron correlation

The energy lowering introduced by adding enough flexibility in the wave function, in order to describe the system qualitatively, is sometimes called the static electron correlation. This is essentially the effect of allowing orbitals to become (partly) singly occupied instead of forcing double occupation, i.e. describing near-degeneracy effects (two or more configurations that almost have the same energy). The remaining energy lowering by correlating the motion of the electrons is called dynamic correlation. The problem is that there is no rigorous way of separating these effects.

Increasing the number of configurations in an MCSCF will recover more and more of the dynamical correlation, until at the full CI limit where the correlation treatment is exact. As mentioned above, MCSCF methods are mainly used for generating a qualitatively correct wave function, i.e. recovering the "static" part of the correlation. However, in order to obtain accurate excitation energies, it is normally necessary to include dynamical correlation. For example, by using the CASPT2 method. Where taking as a reference wave function the CASSCF one, the dynamic correlation is included through the many body perturbation theory.

Møller and Plesset developed the MP perturbation theory.¹¹⁸ MP can systematically approach the exact solution to the Schrödinger equation⁹⁴ order by order of the expansion. The zeroth order of expansion is the HF solution. Thus, MP2 is the second order correction to the HF solution, MP3

is the third order correction, and so on. The order of the expansion rapidly increases the computational cost of the calculation. Experience shows that usually MP3 does not improve the MP2 results, consequently the second order expansion MP2 is the most used.

2.4.1 CASPT2

The Møller-Plesset perturbation theory can be applied also using a CASSCF function as a reference wave function, this is called CASPT2.^{119,120} This method is one of the most accurate *ab initio* approaches and provides a rigorous description of systems with a static electron correlation. The frontier between dynamic and static correlation is unclear. However, most researchers agree that a static electron correlation generally involves few orbitals close in energy to the HOMO. These orbitals have significant fractional occupation numbers, above 0.02 but below 1.98. Usually these orbitals are delocalized and have a strong bonding or antibonding character.

CASPT2 has been applied to study a large variety of systems, specially those related to electronic spectroscopy and photophysics. In this thesis, CASPT2 has been required to properly describe the guanine-indole system. Indole radical cation has the ground and the locally excited states close in energy. This makes that, in a hole transfer reaction, these two states of indole can be implied in the charge transfer reaction.

For a multi-referential case, the wave function is not necessarily an eigenfunction of the n-electron Fock operator (\hat{F}), thus \hat{F} is not a proper zeroth-order Hamiltonian. To restore this, the zeroth-order hamiltonian \hat{H}_0 is defined as:

$$\hat{H}_0 = \hat{P}_0 \hat{F} \hat{P}_0 + \hat{P}_X \hat{F} \hat{P}_X \quad (2.40)$$

where $\hat{P}_0 = |\Psi_0\rangle\langle\Psi_0|$ is a projection operator onto the reference function, \hat{P}_X is the corresponding projection for the rest of configurations, and \hat{F} is the Fock operator.

For the restricted case which assumes closed-shell orbitals and single-determinant wave functions, the Fock operator ($\hat{F}(i)$) for the i-th electron is given by:

$$\hat{F}(i) = \hat{h}(i) + \sum_{j=i}^n [2\hat{J}_j(i) - \hat{K}_j(i)] \quad (2.41)$$

where,

$\hat{h}(i)$ is the one-electron hamiltonian for the i -th electron, n is the total number of occupied orbitals in the system (equal to $[N/2]$, where N is the number of electrons).

$\hat{J}_j(i)$ is the Coulomb operator, defining the repulsive force between the j -th and i -th electrons in the system.

$\hat{K}_j(i)$ is the exchange operator, defining the quantic effect produced by exchanging two electrons.

In the MCSCF matrix, \hat{F} is defined in a way that its diagonal elements correspond to the orbital energies, both for inactive and virtual orbitals. \hat{F} can be written as a one-electron operator:

$$\hat{F} = \sum_{p,q} f_{pq} \hat{E}_{pq} \quad (2.42)$$

Within the CASPT2 method reference wave function, a CAS function is expanded inside a diagonal configurational space. This configurational space is divided in four sets:

\mathbf{V}_0 is a one dimensional space spanned by the CASSCF reference function of the considered state.

\mathbf{V}_K is the space spanned by the orthogonal complement to the reference wave function, in the restricted full CI subspace, used to build the CAS wave function.

\mathbf{V}_{SD} is the space spanned by all single and double excitations of the reference wave function.

\mathbf{V}_X is the rest of the CI space.

The zeroth-order Hamiltonian can be written as:

$$\hat{H}_0 = \hat{P}_0 \hat{F} \hat{P}_0 + \hat{P}_X \hat{F} \hat{P}_X + \hat{P}_K \hat{F} \hat{P}_K + \hat{P}_{SD} \hat{F} \hat{P}_{SD} \quad (2.43)$$

where \hat{P}_0 is the projector onto V_0 , \hat{P}_X is the projector onto V_X ... and \hat{F} is the one particle operator.

Notice that only vectors belonging to V_{SD} contributes to the expansion of the first order wave function. Hence, only those configurations that interact directly with the CAS reference wave function have to be included in the first order wave function, thus only SD are included.

2.4.2 MS-CASPT2

In CASPT2 calculations problems may appear when the CASSCF reference wave function is not a good zero order reference wave function. The CASSCF states, which are orthogonal to the reference state, are part of the secondary space. However, these states do not interact with the reference wave function via the total Hamiltonian. Strong mixtures between the reference CASSCF state and one or more states of the secondary space may occur. To solve these errors, the Multi State CASPT2 (MS-CASPT2) approach can be used.¹²¹

MS-CASPT2 uses a multidimensional reference space. It is computed with a perturbational approach of effective Hamiltonian. This effective Hamiltonian allows the CASSCF states to interact via non diagonal terms. Notice that the diagonal terms are the single state energies (SS-CASPT2).

2.5 Structures

(Guanine-Indole) \bullet^+ systems are the most studied systems in this thesis. We have used the Indole moiety of Tryptophan in the theoretical calculations, instead of the whole Trp, in order to optimize the computational cost of the calculations.

2.5.1 Structural parameters

In the following chapters, in order to define the mutual position between the two monomers of the studied complexes, six parameters have been employed. Three translational ones (shift, slide and rise) and three rotational ones (roll, tilt and twist).^{122,123} In Figure 2.3, these movements are represented.

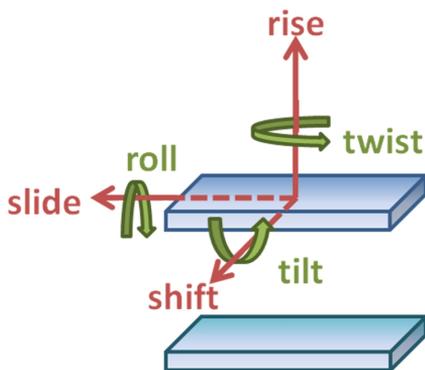


Figure 2.3: Structural parameters that define the mutual arrangements of 2 moieties forming a complex.

Chapter 3

Objectives

The aim of this work is to gain a better understanding in nucleobase-amino acid systems and their hole transfer abilities. NB-*aaa* systems stand for a representative simplification of DNA-protein interactions, which have a crucial role in biological systems repairing oxidative damage on DNA sequence.

Being the computational study of hole transfer reactions in DNA-protein systems a relatively new research area, many questions have yet to be answered. The results presented in this thesis are linear in time, and the knowledge obtained in each paper has been used in the following ones.

The objectives can be divided according to articles and studies presented in this thesis.

Chapter 4 *"Electron transfer from aromatic amino acids to guanine and adenine radical cations in π -stacked and T-shaped complexes"*.

- Determine if π -stacked interactions are required to obtain strong electronic couplings in these systems.
- Calculate the order of the HT rates of NB-aromatic amino acids interactions in order to determine if these interactions could be favorable to extract cationic charges from DNA.

Chapter 5 *"Conformational Dependence of the Electronic Coupling in Guanine-Tryptophan Complexes: A DFT Study"*.

- Analyze the dependence of electronic coupling (which mainly determine the HT rate) on conformational parameters between the subunits of the system.

- Compare the electronic coupling behavior of these NB-aromatic amino acids systems with the one presented by NB-NB interactions within DNA.

Chapter 6 *"Conformational dependence of the electronic coupling for hole transfer between adenine and tryptophan".*

- Adenine can also act as hole acceptor in oxidized DNA when the bridges between guanines are large. Moreover in chapter 4 high charge transfer rates have been observed for adenine-aromatic amino acids interactions. The main objective of this chapter will be to determine the conformational dependence of electronic coupling on (A-Trp)^{•+} system.

Chapter 7 *"Hole Transfer in Guanine-Indole Systems: A Multi-Configurational Study".*

- The indole radical cation has the first 2 states extremely close in energy. The influence of the second state of indole on (G-Ind)^{•+} HT reactions will be analyzed.
- Check if the use of a 3-state method to calculate the electronic coupling could give a better description than the direct method which is a 2-state method usually employed in biological systems.

Chapter 8 *"Hole Transfer in Guanine-Indole Systems: Comparison of electronic coupling obtained with DFT and MS-CASPT2 levels of theory".*

- Using as benchmark data electronic coupling values obtained with the *direct* and the 3-state GMH methods with a MS-CASPT2 level of theory, HF and several DFT functionals (specially long-range corrected ones: CAM-B3LYP, LC-wPBE, and wB97X-D) are tested.

Chapter 4

Electron transfer from aromatic amino acids to guanine and adenine radical cations in π -stacked and T-shaped complexes

Butchosa, C.; Simon, S.; Voityuk, A. A.

Org. Biomol. Chem. **2010**, 8, 1870-1875.

Electron transfer from aromatic amino acids to guanine and adenine radical cations in π stacked and T-shaped complexes†

Cristina Butchosa,^a Sílvia Simon^{*a} and Alexander A. Voityuk^{*a,b}

Received 4th January 2010, Accepted 9th February 2010

First published as an Advance Article on the web 25th February 2010

DOI: 10.1039/b927134a

Similar redox properties of the natural nucleobases and aromatic amino acids make it possible for electron transfer (ET) to occur between these sites in protein–nucleic acid complexes. Using DFT calculations, we estimate the ET rate from aromatic amino acid X (X = Phe, His, Tyr and Trp) to radical cations of guanine (G) and adenine (A) in dimers G–X and A–X with different arrangement of the subunits. We show that irrespective of the mutual orientation of the aromatic rings, the electronic interaction in the systems is strong enough to ensure effective ET from X to G⁺ or A⁺. Surprisingly, relatively high ET rates are found in T-shaped dimers. This suggests that π stacking of nucleobases and aromatic amino acids is not required for feasible ET. In most complexes [G–X]⁺ and [A–X]⁺, we find the excess charge to be confined to a single site, either the nucleobase or amino acid X. Then, conformational changes may initiate migration of the radical cation state from the nucleobase to X and back. The ET process from Trp and Tyr to G⁺ is found to be faster than deprotonation of G⁺. Because the last reaction may lead to the formation of highly mutagenic species, the efficient repair of G⁺ may play an important role in the protection of genomic DNA from oxidative damage.

Introduction

Free radicals as well as X-ray and γ -irradiation are known to generate radical cation and radical anion states of the natural nucleobases in DNA. These states are precursors of highly reactive and mutagenic species that may cause essential damage to DNA producing chemically modified nucleobases, single and double strand breaks, protein–DNA cross-links *etc.*^{1–10} As DNA is an efficient carrier of hole^{11,12} and excess electron charges,^{13–16} the generated charge may migrate through the π -stack long distances away from the site of its initial formation and then initiate a DNA lesion. In living cells, this ability of DNA can be employed for redox sensing and signaling in the genome.¹⁷ Several experiments *in vitro* have demonstrated that radical cation states in DNA are transmitted over distances up to ~ 200 Å.^{18,19} In the past decade, long-distance charge transfer mediated by DNA has received considerable experimental^{11,12} and theoretical attention.^{20,21}

Theoretical methods provide a variety of quantities that are difficult to obtain experimentally and allow one to consider in detail different factors that control the charge transfer process (see recent studies^{22–32} and references therein). Although the main aspects of electron transfer (ET) in DNA are now well understood *in vitro*, many important mechanistic details on ET in genomic DNA remain to be explored. It has been experimentally found that protein–nucleic acid interactions in nucleosome core particles (NCP) can considerably influence the ET process,^{33–35} and therefore, theoretical studies of related models are now of special interest. Using a relatively simple quantum mechanical approach,

Koslowski and coworkers studied the migration of a radical cation through DNA in NCP.³⁶ They suggested that damage to DNA in NCP may occur because of charge transfer from an unprotected DNA segment to the histone-coordinated sequence. Therefore, to protect the genome some mechanisms should exist that prevent the effective hole transfer within the DNA stack. The G⁺ state can undergo rapid deprotonation generating a neutral radical G(H)[•]. The repair of this species implies both electron and proton transfer reactions. This mechanism has been recently studied in detail by density functional theory.³⁷ In the paper, we consider the repair of G⁺ by electron transfer from aromatic amino acids.

The removal of a single electron from a nucleobase results in the formation of an electron deficient site, or hole. A hole generated in DNA is expected to quickly localize at the nearest purine bases, guanine or adenine, which have lower oxidation potentials than pyrimidine bases.¹¹ Thus, the radical cation state G⁺ or (less probable) A⁺ is generated. As the rate of irreversible trapping of the hole due to its chemical reaction with water, oxygen and other species is relatively slow,^{11,12} the hole migrates within DNA using G and A nucleobases as stepping stones.³⁸ In protein–DNA complexes, an amino acid residues X that has a lower oxidation potential than G and A, can act as electron donor (or, equivalently, hole acceptor) retrieving the native state of a nucleobase N from its radical cation¹⁷



This ET reaction should prevent a possible damage to DNA. The low oxidation potentials of tryptophan (Trp) and tyrosine (Tyr) make the repair of G⁺ and A⁺ feasible as has been observed for different systems in aqueous solution,^{39–41} DNA–tripeptide^{42–44} and DNA–protein complexes.^{34,35,45,46} In particular, charge migration in DNA is shown to decrease remarkably by its binding by endonuclease.⁴⁵ Significant differences in the

^aDepartament de Química and Institut de Química Computacional, Universitat de Girona, 17071, Girona, Spain. E-mail: silvia.simon@udg.edu

^bInstitució Catalana de Recerca i Estudis Avançats, Barcelona, Spain. E-mail: alexander.voityuk@icrea.es

† Electronic supplementary information (ESI) available: Additional figures and tables. See DOI: 10.1039/b927134a

dynamics of DNA-mediated hole transport in the presence and absence of packaging into NCP have been reported.^{35,34} In NCP, there are numerous close contacts between DNA and amino acid residues,⁴⁷ which should make possible the electron transfer reaction from X to N⁺. We note that electrostatic interactions between nucleobases, and surrounding amino acid residues and water molecules, affect the stability of G⁺ and A⁺.⁴⁸ Thus, the standard oxidation potentials of N and X provide only rough estimates for the ET free energy.

The hole trapping process can be accompanied by proton transfer. The formation of radical cation X⁺ leads to a decrease of its pK_a-value and can enforce rapid deprotonation of the residue due to proton transfer to surroundings. As a result, back ET from G to X becomes unfeasible. A general theoretical approach for treatment of proton-coupled electron transfer reactions is described by Hammes-Schiffer.⁴⁹

The direct repair of N⁺ (eqn (1)) will play an important role if the rate of this process is comparable or higher than that of competitive reactions. According to Marcus equation,⁵⁰ three parameters (the electronic coupling V , the driving force ΔE and the reorganization energy λ) determine the ET rate at the temperature T :

$$k_{ET} = \frac{2\pi}{h} V^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(\frac{-(\Delta E + \lambda)^2}{4\lambda k_B T}\right) \quad (2)$$

The dependence of k_{ET} on the mutual orientation of donor and acceptor is mainly controlled by V , which is crudely proportional to the overlap of the orbitals of donor and acceptor. The driving force ΔE is the difference of redox potentials of the donor and acceptor. ΔE for charge shifting $DA^+ \rightarrow D^+A$ is almost independent of the distance between the donor and acceptor; however, when the redox sites possess dipole moments, ΔE may remarkably change by mutual rotation of D and A as demonstrated below. The reorganization energy λ is the change in the driving force to move the reactants to the product configuration without actually transferring the electron. For ET in biomolecules, the reorganization energy λ is usually assumed to be in the range 0.5 to 1.5 eV. In our estimation, $\lambda = 1.0$ eV is employed.

In the present study, we calculate the ET rate for several model systems N–X, where N is a purine nucleobase (N = G and A) and X is a truncated aromatic amino acid (Trp, Tyr, His or Phe), and consider its dependence upon the nature of N and X and their mutual orientation in the dimers. Our starting point is the stacked and T-shaped structures of G–X and A–X recently reported by Wetmore and co-workers.⁵¹ The potential energy surface of these complexes was systematically studied at the MP2/6-31* level of theory; the stabilization energies were calculated using the CCSD(T)/CBS method. On the basis of the high-level calculations, Wetmore *et al.* concluded that (1) both stacking and T-shaped interactions are very close in magnitude to biologically relevant hydrogen bonds and (2) the interaction of monomers in T-shaped dimers is as strong as their stacking interaction.⁵¹ It means that T-shaped conformations may play an important role in protein–DNA complexes. For all these dimers, we carry out DFT calculations of V and ΔE , and estimate the ET rates in the systems. We obtain that the probability of ET in complexes depends critically on the nature of N and X as well as on the dimer structure. Depending on the mutual orientation

of the subunits in complexes, the radical cations G⁺ and A⁺ are stabilized or destabilized as compared with their isolated states. Interestingly, the electron hole localized at G can migrate to Trp and back when passing from one dimer conformation to another. The relatively high ET rates we have found for T-shaped complexes suggest that π stacking of nucleobases and aromatic amino acids is not required for feasible ET from X to N⁺.

Computational details

Structure

All optimized structures found by Wetmore *et al.*⁵¹ for G–X and A–X dimers were studied. Besides two stacked (S1 and S2) structures, we considered tree T-shaped conformation (E, F1 and F2), where the plane of N is almost perpendicular to that of X. The structures of G–Trp are shown in Fig. 1. For the sake of simplicity, we used a slightly different notation for the conformers. Its correspondence to the original notation⁵¹ is explained in the ESI.†

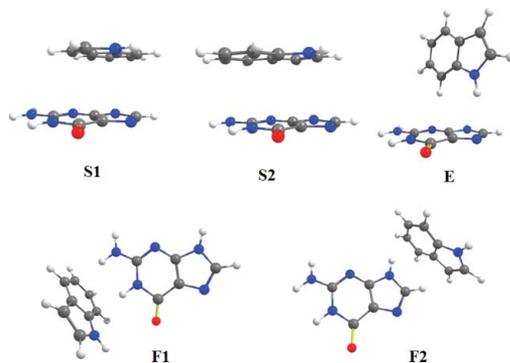
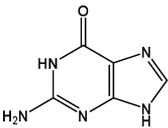
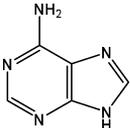
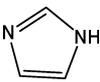
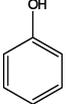
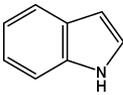


Fig. 1 Structure of the guanine-tryptophan complexes: stacked conformations S1 and S2, and T-shaped conformations E (edge) and F1 and F2 (face).

ET parameters

It has been shown that reliable estimates for the driving force ΔE and electronic coupling V can be obtained using Kohn–Sham orbitals stemming from DFT calculations of neutral systems.^{52,53} In particular, the B3LYP results for radical-cation states of nucleobase dimers are in good agreement with the data of high-level calculations (CASSCF and CASPT2).⁵⁴ The diabatic energies and electronic couplings of donor and acceptor were calculated within the two-state model. First, we compared three methods: the Generalized Mulliken–Hush,⁵⁵ Fragment Charge Difference⁵⁶ and the Direct method.^{22,25,57} In most cases, these schemes provide very similar results (see the ESI†). The direct method is computationally very robust, and it has been successfully used in DFT calculations of the ET parameters in DNA.²⁵ So, this scheme is employed throughout our study. The DFT calculations were carried out using the standard B3LYP functional,⁵⁸ and 6-31G* basis set. We employed the program *Gaussian 03*.⁵⁹

Table 1 The structure of monomers and the relative energy of the radical cation states ϵ_{rel} (as compared to G^+)

Monomer	Structure	$\epsilon_{\text{rel}}/\text{eV}$
G		0.00
A		0.367
His		0.583
Phe		1.192
Tyr		0.462
Trp		-0.125

Results and discussion

Isolated fragments

Relative values of the ionization energy of monomers N and X (Table 1) provide preliminary estimates of the ET energy for $N^+ + X \rightarrow N + X^+$. As shown, Trp is the strongest electron donor. Its ionization energy is even lower than that of G. Then, Tyr and His may be involved in the repair of A^+ , while ET from Phe to G^+ and A^+ is hardly possible. Interestingly, the ionization energy of A is 0.37 eV higher than that of G in line with the experimental oxidation potentials, 1.7 and 1.3 V.⁶⁰

Stabilization/destabilization of N^+ and X^+ in dimers

When monomers N and X form a complex, their ionization energies change. Because of the interaction within the dimer, the states N^+ and X^+ may be stabilized or destabilized. Obviously, the interaction energy depends on the nature and the arrangement of monomers. Fig. 2 shows how the energy of the radical-cation states in G–Trp and A–Trp depends on the dimer conformation. The data for the other complexes are listed in the ESI.† In G–Trp, G^+ is stabilized in stacked (S1, S2) and two T-shaped (F1, F2) structures (see Fig. 1). The stabilization energy Δ is ~ 0.2 – 0.3 eV. In contrast, G^+ is significantly destabilized in

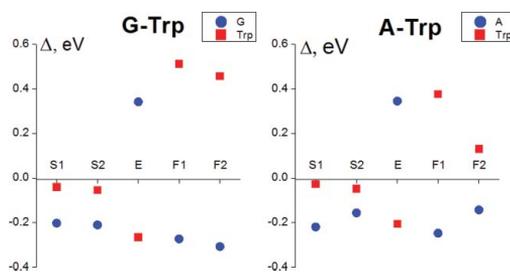


Fig. 2 Stabilization/destabilization energy of radical cation states localized on the nucleobase or on Trp in G–Trp and A–Trp dimers.

the edge (E) conformer. Quite different changes are found for Trp^+ . Its energy remains almost unchanged (as compared with the isolated radical cation) in S1 and S2, while the Trp^+ state is significantly stabilized in E and destabilized in F1 and F2. Qualitatively similar changes are found in the A–Trp dimers (Fig. 2). We note that the Δ values can be estimated using a simple electrostatic model (the ion–dipole interaction energy). Because the dipole moment of A, $\mu(A)$, is smaller than $\mu(G)$, less significant stabilization/destabilization energies for X^+ are found in A–Trp. Since in E and F conformations, the dipole moments of the monomers are of opposite directions, the Δ values for N^+ and X^+ change their sign by passing from E to F1 and F2 (see Fig. 2).

ET energy

Fig. 3 shows the calculated values of the driving force ΔE for ET from X to N^+ in the dimers. In most G–X structures, ΔE is positive, and therefore, the ET process is unlikely. Negative ΔE values are found in the E conformation of G–X, where X = Trp, Tyr and His. As the ionization energy of A by 0.4 eV larger than that of G, the A^+ state can be reduced more easily (in Fig. 3, the ΔE values found for A–X are more negative than for G–X). Independent of the conformation of G–Phe and A–Phe, the ET driving force is calculated to be positive in these complexes. As expected, Trp is the best reducing agent among the aromatic amino acids. Tyr and His have very similar ionization energies.

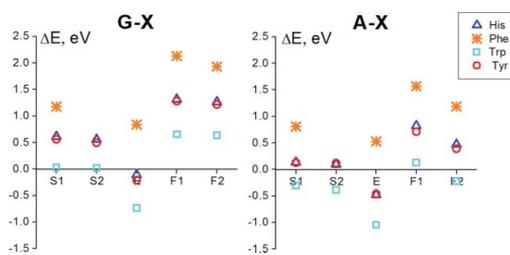


Fig. 3 Dependence of the ET driving force on the dimer conformation.

Electronic couplings

Computed values of the electronic coupling V in dimers are shown in Fig. 4. As seen, the coupling is strongly dependent on the mutual arrangement of monomers. The ET rate is proportional

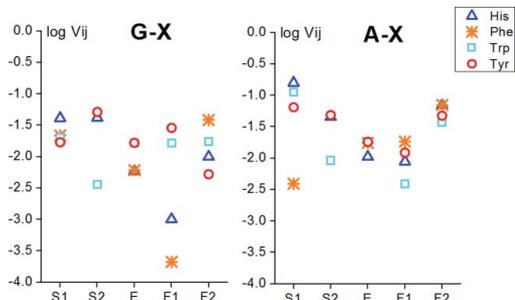


Fig. 4 Dependence of the electronic coupling on the dimer conformation.

to V^2 , eqn (2), and therefore, it should be much more sensitive to conformational changes. Because no general trend is observed for the complexes, the conformational dependence of V is difficult to predict without quantum chemical calculations. As seen from Fig. 4, $V(\text{G-His})$ remains almost unperturbed by passing from S1 to S2, while there are remarkable differences in other complexes. Large variations of V are found in A-X. Surprisingly, the coupling values calculated in T-shaped structures (E, F1 and F2) are comparable in magnitude with those in stacked dimers. Therefore, π stacking of nucleobases and aromatic amino acids is not required for feasible ET between these sites. As has been already demonstrated for natural and modified DNA, small conformational changes may drastically affect the electronic coupling^{21,27-32} and averaging over many conformation should be applied to get accurate estimates of the effective coupling. We note that the averaging over thermally available structures considerably decreases the extent to which the “observed” ET rate is actually dependent on conformational changes.⁶¹

Excess charge distribution

Let us consider now the excess charge localization in the ground state of radical cations G-X and A-X. Fig. 5 displays a charge difference $\Delta Q = Q(N) - Q(X)$ in the dimers. $\Delta Q = 1$ means that the excess charge is completely localized on the nucleobase; if $\Delta Q = -1$, the radical cation state is localized on X; when the excess charge is delocalized over the system, $|\Delta Q|$ is close to zero. There is a simple relation between ΔQ and the ET parameters ΔE and V ⁶²

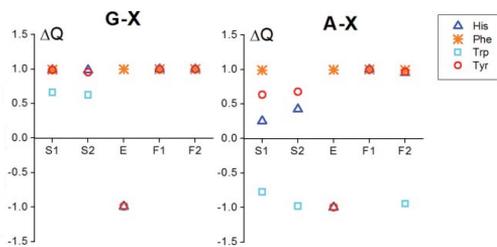


Fig. 5 Difference in charges on the nucleobase and residue X, $\Delta Q = Q(N) - Q(X)$ in radical cations G-X and A-X.

$$\Delta Q = \frac{\Delta E}{\sqrt{\Delta E^2 + 4V^2}} \quad (3)$$

Because in most complexes G-X, ΔE is positive and the coupling is relatively small, the excess charge is mainly confined to G. Only in the edge dimers with $X = \text{Trp}$, Tyr and His, where $\Delta E < 0$, the radical cation state is found on X. Since in the stacked dimers G-Trp, absolute values of ΔE and V are similar, some delocalization of the excess charge is found. In A-X dimers, the charge distribution is different. Irrespective of the mutual position of A and Trp, the radical cation state is localized on Trp. The charge is delocalized in the stacked dimers A-His and A-Tyr. Overall in line with eqn (3), ΔE and ΔQ show similar trends (see Fig. 3 and 5). We note that in spite of low activation energies required for conformational transitions in the dimers, considerable redistribution of the charge and spin density may be induced when passing from one conformation to another.

ET rates

Using the calculated values of ΔE and V , and the reorganization energy $\lambda = 1$ eV, we estimated the rate of electron transfer $N^+ + X \rightarrow N + X^+$ in the dimers. Fig. 6 shows the ET rates faster than 10^6 s^{-1} . We remind that eqn (2) can be applied only to systems with weak coupling (nonadiabatic regime). Because of that, the ET rate for dimers with $V > 0.06$ eV was not calculated. As seen from Fig. 6, in several dimers G-X and A-X, the ET rate is found to be $\sim 10^8 \text{ s}^{-1}$ or higher. It has been experimentally found that G^+ in DNA deprotonates quite rapidly (10^7 s^{-1}), forming guanyl radical $[G(-H)]$.⁶³ The last species is very reactive and may lead to mutagenic effects. Because the ET from Trp, Tyr and His to G^+ is found to be faster than deprotonation of G^+ , it may be important for protecting genome DNA.¹⁷ The values of the absolute ET rate depend on the parameter λ . If the ET driving force is close to zero (e.g. in G-Trp dimers) the temperature dependence of the rate is approximately determined by $\exp(-\lambda/4kT)$. At room temperature, a variation δ of the reorganization energy will be translated in a factor of $10^{-4.2\delta}$ (δ in eV). Thus, using $\lambda = 0.8$ eV instead of $\lambda = 1.0$ eV ($\delta = -0.2$ eV) one increases the ET rate by a factor of 7; to the contrary, the estimated rate will decrease by the same factor when $\lambda = 1.2$ eV is employed in eqn (2). Obviously, relative values

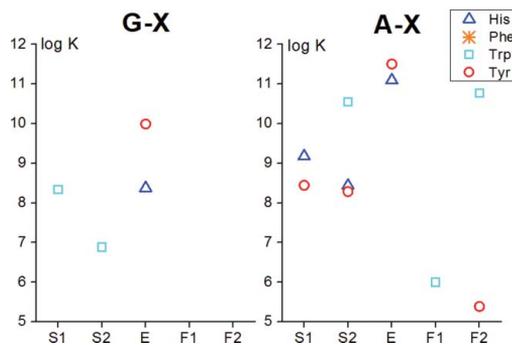


Fig. 6 Dependence of the ET rate on the dimer conformation in G-X (at the left) and A-X (at the right).

of the ET rate calculated for different dimer conformations are much less sensitive to the parameter λ .

The results obtained agree well with available experimental data. In particular, it has been observed that both the Tyr radical and the Trp radical can be generated in DNA–tripeptide complexes by ET from these residues to G^+ .^{42,43} As Fig. 6 suggests, effective ET should occur in both G–Trp (stacked complexes) and G–Tyr (edge complex).

Conclusions

The efficient ET process between amino acid residues and guanine (or adenine) radical cations (G^+ or A^+) may play an important role in protection of genomic DNA from oxidative damage.¹⁷ Not much is still known; however, about ET in DNA–protein complexes. In this paper, we have studied how the mutual arrangement of nucleobases and aromatic amino acid residues X can affect the rate of ET between these species. Using the optimized structures found by Wetmore *et al.*⁵¹ for stacked and T-shaped dimers G–X and A–X, we carried out DFT calculations of the ET parameters (the driving force and electronic coupling) and estimated the ET rate in the complexes. The following results have been obtained.

(1) Irrespective of the orientation of subunits within the system, the electronic couplings are strong enough to ensure effective charge transfer from aromatic amino acid residues to G^+ or A^+ . While quite strong electronic interaction between π -stacked molecules is usually expected, the relatively large coupling values found in T-shaped dimers, where the aromatic rings of subunits are perpendicular to each other, are quite surprising. This finding clearly shows that for efficient ET in DNA–protein complexes, π stacking of nucleobases and aromatic amino acids is not required.

(2) In the dimers, the driving force of ET is shown to be strongly dependent on the mutual orientation of the monomers. The most favourable values are calculated for edge configurations.

(3) In most N–X complexes, the excess charge and spin density is confined to a single subunit, either to nucleobase N or amino acid residue X. Changes in the monomer orientation may lead to migration of the radical cation state between the N and X sites.

(4) ET transfer from Trp to G^+ is found to be faster than deprotonation G^+ , which can be followed the formation of highly mutagenic species. Thus, the ET reaction $[N^+ X] \rightarrow [N X^+]$ may play an important role in protection of genomic DNA from oxidative damage.

Obviously, for a more realistic description of ET from amino acid residues to radical cation states of nucleobases, the effects of structural fluctuations and interactions with the environment should be taken into account. However, we believe that the main results obtained in this study are applicable to extended DNA–protein systems.

Acknowledgements

The authors greatly acknowledge Dr S. D. Wetmore for providing atomic coordinates for the dimers considered in the paper. The work was supported by the Spanish *Ministerio de Educación y Ciencia* (Project No CTQ2009-12346) and by the Generalitat de Catalunya (Project No. FI-DGR2009 modalitat B).

References

- 1 C. J. Burrows and J. G. Muller, *Chem. Rev.*, 1998, **98**, 1109–1151.
- 2 A. Kumar and M. D. Sevilla, *J. Am. Chem. Soc.*, 2008, **130**, 2130–2131.
- 3 P. Swiderek, *Angew. Chem., Int. Ed.*, 2006, **45**, 4056.
- 4 S. Perrier, J. Hau, D. Gasparutto, J. Cadet, A. Favier and J. L. Ravanat, *J. Am. Chem. Soc.*, 2006, **128**, 5703–5710.
- 5 X. Xu, J. G. Muller, Y. Ye and C. J. Burrows, *J. Am. Chem. Soc.*, 2008, **130**, 703–709.
- 6 A. Kupan, A. Saulier, S. Broussy, C. Seguy, G. Pratiel and B. Meunier, *ChemBioChem*, 2006, **7**, 125–133.
- 7 J. Llano and L. A. Eriksson, *Phys. Chem. Chem. Phys.*, 2004, **6**, 4707–4713.
- 8 W. Luo, J. G. Muller and C. J. Burrows, *Org. Lett.*, 2001, **3**, 2801–2804.
- 9 R. Misiaszek, C. Cream, A. Joffe and N. E. Geacintov, *J. Biol. Chem.*, 2004, **279**, 32106–32115.
- 10 L. I. Shukla, A. Adhikary, R. Pazdro, D. Becker and M. D. Sevilla, *Nucleic Acids Res.*, 2004, **32**, 6565–6574.
- 11 G. B. Schuster, Long-Range Charge Transfer in DNA I-II, *Topics in Current Chemistry*, 2004.
- 12 J. C. Genereux and J. K. Barton, *Chem. Rev.*, 2010, **110**, DOI: 10.1021/cr900228f.
- 13 H.-A. Wagenknecht, *Angew. Chem., Int. Ed.*, 2003, **42**, 2454–2460.
- 14 C. Behrens, L. T. Burgdorf, A. Schwöglar and T. Carell, *Angew. Chem., Int. Ed.*, 2002, **41**, 1763.
- 15 T. Ito and S. E. Rokita, *J. Am. Chem. Soc.*, 2004, **126**, 15552–15559.
- 16 P. Daublain, A. K. Thazhathveetil, Q. Wang, A. Trifonov, T. Fiebig and F. D. Lewis, *J. Am. Chem. Soc.*, 2009, **131**, 16790–16797.
- 17 J. C. Genereux, A. K. Boal and J. K. Barton, *J. Am. Chem. Soc.*, 2010, **132**, 891–905.
- 18 M. E. Nunez, D. B. Hall and J. K. Barton, *Chem. Biol.*, 1999, **6**, 85–97.
- 19 P. T. Henderson, D. Jones, G. Hampikian, Y. Kan and G. B. Schuster, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 8353.
- 20 Y. A. Berlin, I. V. Kurnikov, D. N. Bertran, M. A. Ratner and A. L. Burin, *Top. Curr. Chem.*, 2004, **237**, 1–35.
- 21 A. A. Voityuk, Computational modeling of charge transfer in DNA, in *Computational studies of RNA and DNA*, ed. J. Šponer and F. Lankas, Springer, Dordrecht, 2006, pp. 485–512.
- 22 A. Troisi and G. Orlandi, *J. Phys. Chem. B*, 2002, **106**, 2093–2101.
- 23 A. A. Voityuk, K. Siriwoong and N. Rösch, *Angew. Chem., Int. Ed.*, 2004, **43**, 624–627.
- 24 N. Rösch and A. A. Voityuk, *Top. Curr. Chem.*, 2004, **237**, 37–72.
- 25 K. Senthilkumar, F. C. Grozema, C. F. Guerra, F. M. Bickelhaupt, F. D. Lewis, Y. A. Berlin, M. A. Ratner and L. D. A. Siebbeles, *J. Am. Chem. Soc.*, 2005, **127**, 14894.
- 26 F. C. Grozema, S. Tonzani, Y. A. Berlin, G. C. Schatz, L. D. A. Siebbeles and M. A. Ratner, *J. Am. Chem. Soc.*, 2008, **130**, 5157.
- 27 K. Siriwoong and A. A. Voityuk, *J. Phys. Chem. B*, 2008, **112**, 8181.
- 28 T. Steinbrecher, T. Koslowski and D. A. Case, *J. Phys. Chem. B*, 2008, **112**, 16935.
- 29 D. Reha, W. Barford and S. Harris, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5436.
- 30 T. Kubar and M. V. Elstner, *J. Phys. Chem. B*, 2008, **112**, 8788.
- 31 A. A. Voityuk, *J. Chem. Phys.*, 2008, **128**, 115101.
- 32 W. He, E. Hatcher, A. Balaeff, D. N. Beratan, R. R. Gil, M. Madrid and C. Achim, *J. Am. Chem. Soc.*, 2008, **130**, 13264.
- 33 S. R. Rajski and J. K. Barton, *Biochemistry*, 2001, **40**, 5556.
- 34 M. E. Nunez, K. T. Noyes and J. K. Barton, *Chem. Biol.*, 2002, **9**, 403.
- 35 C. C. Bjorklund and W. B. Davis, *Nucleic Acids Res.*, 2006, **34**, 1836.
- 36 T. Cramer, S. Krapf and T. Koslowski, *Phys. Chem. Chem. Phys.*, 2004, **6**, 3160.
- 37 N. R. Jena, P. C. Mishra and S. Suhai, *J. Phys. Chem. B*, 2009, **113**, 5633–5644.
- 38 B. Giese, J. Amaudrut, A. K. Kohler, M. Spormann and S. Wessely, *Nature*, 2001, **412**, 318–320.
- 39 P. M. Cullis, G. D. D. Jones, M. C. R. Symons and J. S. Lea, *Nature*, 1987, **330**, 773–774.
- 40 J. R. Milligan, J. A. Aguilera, A. Ly, N. Q. Tran, O. Hoang and J. F. Ward, *Nucleic Acids Res.*, 2003, **31**, 6258–6263.
- 41 J. Pan, W. Lin, W. Wang, Z. Han, C. Lu, S. Yao, N. Lin and D. Zhu, *Biophys. Chem.*, 2001, **89**, 193–199.
- 42 H. A. Wagenknecht, E. D. A. Stemp and J. K. Barton, *Biochemistry*, 2000, **39**, 5483–5491.
- 43 H.-A. Wagenknecht, E. D. A. Stemp and J. K. Barton, *J. Am. Chem. Soc.*, 2000, **122**, 1–7.

-
- 44 E. Mayer-Enthart, P. Kaden and H.-A. Wagenknecht, *Biochemistry*, 2005, **44**, 11749–11757.
- 45 K. Nakatani, C. Dohno, A. Ogawa and I. Saito, *Chem. Biol.*, 2002, **9**, 361–366.
- 46 H.-A. Wagenknecht, S. R. Rajski, M. Pascaly, E. D. A. Stemp and J. K. Barton, *J. Am. Chem. Soc.*, 2001, **123**, 4400–4407.
- 47 C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder and T. J. Richmond, *J. Mol. Biol.*, 2002, **319**, 1097–1113.
- 48 A. A. Voityuk and W. B. Davis, *J. Phys. Chem. B*, 2007, **111**, 2976–2985.
- 49 S. Hammes-Schiffer, *Acc. Chem. Res.*, 2009, **42**, 1881–1889.
- 50 R. A. Marcus and N. Sutin, *Biochim. Biophys. Acta*, 1985, **811**, 265–322.
- 51 L. R. Rutledge, H. F. Durst and S. D. Wetmore, *J. Chem. Theory Comput.*, 2009, **5**, 1400–1410.
- 52 M. Félix and A. A. Voityuk, *J. Phys. Chem. A*, 2008, **112**, 9043–9049.
- 53 M. Félix and A. A. Voityuk, *Int. J. Quantum Chem.*, 2010, DOI: 10.1002/qua.22419.
- 54 L. Blancafort and A. A. Voityuk, *J. Phys. Chem. A*, 2006, **110**, 6426–6432.
- 55 R. J. Cave and M. D. Newton, *Chem. Phys. Lett.*, 1996, **249**, 15–19.
- 56 A. A. Voityuk and N. Rösch, *J. Chem. Phys.*, 2002, **117**, 5607–5616.
- 57 M. D. Newton, *Chem. Rev.*, 1991, **91**, 767–792.
- 58 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
- 59 M. J. Frisch, *et al.*, *Gaussian 03, Rev. E.03*, 2004, Gaussian, Inc., Wallingford.
- 60 S. Doose, H. Neuweiler and M. Sauer, *ChemPhysChem*, 2009, **10**, 1389–1398.
- 61 A. A. Voityuk, *J. Phys. Chem. B*, 2009, **113**, 14365–1468.
- 62 A. A. Voityuk, *J. Phys. Chem. B*, 2005, **109**, 10793–10796.
- 63 K. Kobayashi and S. Tagawa, *J. Am. Chem. Soc.*, 2003, **125**, 10213–10218.

Chapter 5

Conformational Dependence of the Electronic Coupling in Guanine-Tryptophan Complexes: A DFT Study

Butchosa, C.; Simon, S.; Voityuk, A. A.

Int. J. Quantum Chem. **2012**, 112, 1838-1843.

Butchosa, C., Simon, S. and Voityuk, A. A. "Conformational dependence of the electronic coupling in guanine–tryptophan complexes: A DFT study". *International journal of quantum chemistry*. Vol. 112, issue 7 (April 5 2012) : p. 1838–1843.

Copyright © 2011 Wiley Periodicals, Inc.

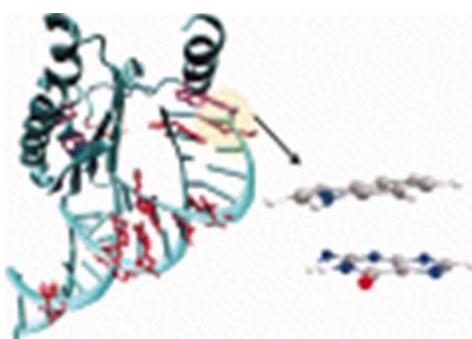
Article first published online: 4 APR 2011

<http://onlinelibrary.wiley.com/doi/10.1002/qua.23077/abstract>

DOI: <http://dx.doi.org/10.1002/qua.23077>

Abstract:

Radical cation states (electron holes) generated in DNA lead to mutagenic effects. In DNA–protein complexes, the holes can migrate from DNA to aromatic amino acid residues preventing the formation of the oxidative lesions. The efficiency of hole transfer (HT) may significantly depend on the arrangement of redox sites. This dependence is mainly determined by sensitivity of the electronic coupling of donor and acceptor to structural changes. Based on DFT calculations of a number of guanine–tryptophan (G–Trp) complexes, we explore the conformational dependence of HT electronic coupling in this dyad. Stacked and T-shaped structures are considered. The electronic coupling in the system is shown to be responsive to the mutual arrangement of G and Trp. Although in most cases the probability of HT in T-shaped conformations is predicted to be lower than in π stacks, several T-type structures are found where HT should be very efficient.



Keywords: electron transfer; electronic coupling; DFT; DNA–protein complexes

Chapter 6

Conformational dependence of the electronic coupling for hole transfer between adenine and tryptophan

Butchosa, C.; Simon, S.; Voityuk, A. A.

Comput. Theoret. Chem. **2011**, 975, 38.

Butchosa, C., Simon, S., Voityuk, A.A. "Conformational dependence of the electronic coupling for hole transfer between adenine and tryptophan". *Computational and Theoretical Chemistry*. Vol 975, issues 1-3 (November 15, 2011) : p. 38-41

Copyright © 2011 Elsevier B.V. All rights reserved

Received 18 November 2010. Revised 13 April 2011. Accepted 19 April 2011. Available online 30 April 2011.

<http://www.sciencedirect.com/science/article/pii/S2210271X11002350>

<http://dx.doi.org/10.1016/j.comptc.2011.04.025>

Abstract

Oxidation of DNA may lead to mutagenic lesions generated far away from the initial oxidized site because of migration of radical cation states through the π stack. In DNA–protein complexes, transfer of the excess positive charge from nucleobases to aromatic amino acid residues protects DNA from possible mutations. In the present paper, we explore how the probability of the hole transfer (HT) process between adenine (A) and tryptophan (Trp) depends on the mutual position of these sites. To accomplish this, we carry out DFT calculations of HT electronic coupling in different conformations of the A-Trp complex. Stacked and T-shaped structures are considered. The HT rate in the system is shown to be very sensitive to the mutual position of the nucleobase and amino acid residue. Interestingly, the strongest coupling is obtained in stacked structures where only one of two rings in each molecule are involved in the π – π interaction and a surprisingly weak coupling is found in the eclipsed conformation of the A-Trp complex with the perfect overlap of the aromatic systems. Although the HT rate derived for T-shaped conformations is in most cases slower than in π stacks, several T-shaped conformations are found where the HT process should be quite efficient.

Keywords

- Electron transfer;
- Electronic couplings;
- DNA–protein complexes;
- DFT

Chapter 7

Hole Transfer in Guanine-Indole Systems: A Multi-Configurational Study

Butchosa, C.; Simon, S.; Blancafort, Ll.; Voityuk, A.

Submitted as:

*MS-CASPT2 Study of Hole Transfer in Guanine-Indole complexes:
Effective 2-State Treatment,*

at Journal of Physical Chemistry

7.1 Abstract

Hole transfer (migration of radical cation states) in DNA-protein complexes has attracted considerable interest because it might play an important role preventing oxidative damage of DNA.

The results of several model complexes consisting of guanine and indole are reported. Different π -stacked structures of the radical cations are considered. Calculations were performed at MS-CASPT2 level of theory. Obtained data demonstrate that, the commonly used 2-state model, is of limited use to deal with indole radical cation, and three adiabatic states should be treated to derive proper electronic couplings for the hole transfer reaction in these complexes. Depending on the conformation, indole acts as a 2-state charge acceptor. Thus, their ground and locally excited states are both oxidizable. In such a case, the second indole state becomes energetically accessible and the charge transfer rate from $G^{\bullet+}$ to indole will increase. As a consequence, a 2-state treatment which neglects the second indole state will underestimate the effective HT rate.

7.2 Introduction

Tryptophan (Trp) aromatic amino acid is well known for its abilities as hole acceptor in charge transfer reactions.³² One example of their importance in biological processes is its role on the photoactivation of photolyase protein.^{33–37,124} Photolyase uses reduced flavin cofactor ($FADH^-$) to repair cyclobutane pyrimidine dimers (CPDs) which are the major UV-induced lesions on DNA. There is a chain of three Trps, well conserved through all known photolyases, which photoactivate the FADH0 reducing it mediating an electron transfer reaction which takes ~ 30 ps by a hopping mechanism.³⁸ Moreover, the indole scaffold of Trp is known to be a successful antioxidant acting as free radical scavenger and broad-spectrum antioxidant in living organisms. Melatonin hormone (MLT) is synthesized by the pineal gland from Trp amino acid.¹²⁵ This hormone has an indole scaffold. MLT not only acts as scavenger from free radicals but also stabilizes the membrane cells making them more resistant to oxidative damage.¹²⁶ Moreover, MLT can cross all the morphophysiological barriers and can protect till lipids in the cell membrane to DNA in the nucleus.^{127,128}

Milligan et al.⁴³ demonstrated that Trp can repair guanyl radicals extracting the cationic charge by an electron transfer reaction, with a rate of $1.0 \times 10^7 \text{ dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$. Guanine has the lowest ionization potential within

the four nucleobases (8.02 eV) and the cationic charges use guanines as stepping stones in their long range HT reactions along DNA.^{6,9,129} Thus, guanyl radical is the most probable hole donor to transfer a cationic charge from an oxidized DNA to the Ind scaffold. Furthermore, in previous studies we already explored which type of G-aromatic amino acid conformations are the best ones to perform a HT reaction, finding that (G-Ind)^{•+} interactions present higher values of electronic coupling and HT rates.^{130–132}

However, the (G-Ind)^{•+} system is particularly challenging because the first excited state of Ind radical cation is very close in energy to its ground state. Depending on the conformation, we found a minimum difference of 0.232 eV between the diabatic states. In previous studies on charge transfer reactions involving Ind, it has been assumed that the widely used 2-state model was a suitable approximation to calculate the electronic coupling.^{20,36,37,130–132} Therefore, taking into account the proximity of Ind^{•+} locally excited state to the ground state, employing a 2-state model is questionable and an extension to a 3-state model may become necessary.^{90,101,104,106} The Ionization Potentials (IP) of the ground and the first excited state of Ind are separated by 0.45 eV (Figure 7.1). Analyzing the IP of charge donor nucleobases (guanine and adenine), it seems that charge transfer processes of A-Ind system could be very challenging because adenine also has its ground and excited states IP close in energy.²⁹

In the present chapter, we treat the HT between G and Ind explicitly as a 3-state case and compare the results with the 2-state approach. Using the 3-state model, the diabatic states and the couplings are obtained from the adiabatic energies with the generalized Mulliken-Hush approach.¹⁰⁰ Because of the complexity of the problem, energies and wave functions of the adiabatic electronic states rather than orbital energies (i.e. one electron approximation) are used, where the MS-CASPT2 method is used to obtain the best possible description of the adiabatic states. In addition, the present MS-CASPT2 results represent a benchmark for future studies on HT between DNA and proteins with computationally less expensive DFT based methods.

7.3 Computational details

The CASSCF, CASPT2, and MS-CASPT2 (multistate formulation of CASPT2 which accounts for the nonorthogonality of the CASPT2 wave function)¹²¹ calculations are carried out for several π stacked (G-Ind)^{•+} structures using the ano-s basis set and the program MOLCAS 7.2.¹³³

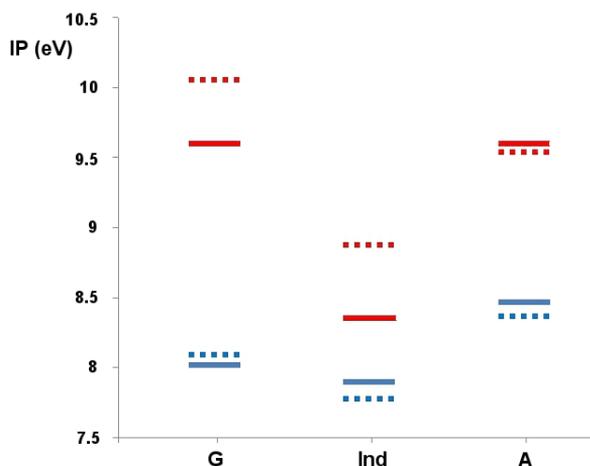


Figure 7.1: Ionization potentials of guanine, indole, and adenine. Both the ionization of the ground state (blue) and of the first excited state (red) are shown. Solid lines are experimental data while dashed lines are calculated at MS-CASPT2 level of theory.²⁹

For $(\text{G-Ind})^{\bullet+}$ calculations, an active space of 11 electrons in 12 orbitals (11,12) has been used. The orbitals were divided equally between the two molecules to describe their π orbitals, which are associated with charge transfer reactions.

The energies of the ground and excited states are obtained from a single point calculation, using state-averaged orbitals over 5 states, with equal weights for each state. The 5 states were computed in order to analyze which ones should be included in the multistate electronic coupling calculation. The dipole moments of each state and the transition dipole moments are derived from the perturbationally modified CAS configuration interaction wave function (PM-CASCI), which is obtained from the MS-CASPT2 calculation. A level shift parameter of 0.2 has been used.

7.3.1 Geometries

A set of six $(\text{G-Ind})^{\bullet+}$ stacked conformations has been studied in this work (Figure 7.2). These structures were obtained from a previous conformational study of electronic coupling in $(\text{G-Ind})^{\bullet+}$ systems.¹³¹ Several structures were selected because they lead to very strong electronic couplings values (π_2 , π_3 , π_4 , π_5). Other structures of interest are also considered, such as π_1 which is a perfect stacking between the two

molecules, and $\pi 6$ which has the Ind interacting with the other face of G.

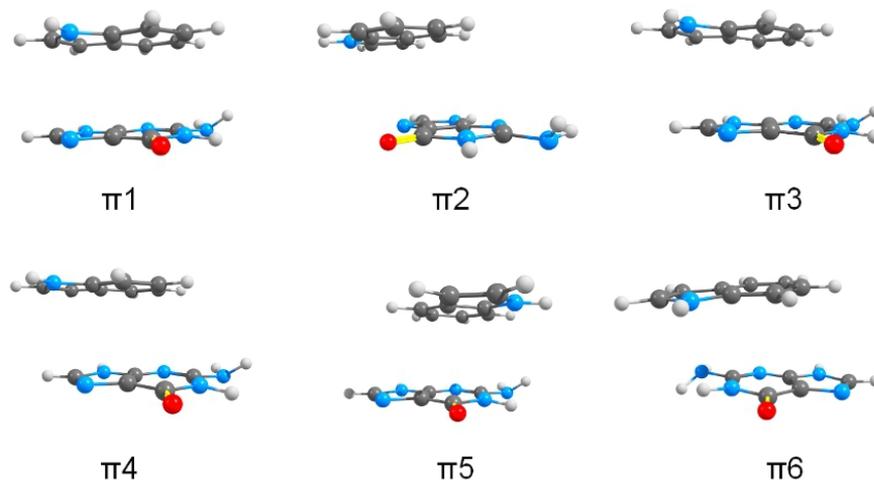


Figure 7.2: $\pi 1$, perfect stack between the aromatic rings of G and Ind which are separated by a rise of 3.38 Å. $\pi 2$, starting from $\pi 1$ a shift movement of 2 Å. $\pi 3$, starting from $\pi 1$ a slide movement of -1.7 Å. $\pi 4$, starting from $\pi 1$ a slide movement of 2.7 Å. $\pi 5$, starting from $\pi 1$ a twist rotation of 110°. $\pi 6$, G-Ind structure where Ind is interacting with the other guanine side.

Notice that small changes in the mutual arrangement of the subunits involved in the HT reactions can lead to significant variations of the coupling between them (the matrix element increases by two orders of magnitude).⁸⁰

7.4 Methodology

The dependence of the HT rate (k_{HT}) on the mutual orientation of the donor and acceptor is mainly determined by the electronic coupling squared, V^2 , as it is pointed in the main equation of Marcus charge transfer theory (Equation 7.1).⁶⁶

$$k_{HT} = \frac{2\pi}{\hbar} |V_{da}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(\frac{-(\Delta G^0 + \lambda)^2}{4\lambda k_B T}\right) \quad (7.1)$$

In Equation 7.1, ΔG^0 is the driving force and λ is the reorganization energy. λ is composed by the internal (λ_{int}) and the solvent (λ_s) terms, $\lambda = \lambda_{int} + \lambda_s$. A temperature of 298.15K and a reorganization energy value of 1 eV have

been used. For charge transfer in biomolecules, the reorganization energy λ is usually assumed to be in the range of 0.5 to 1.5 eV.

To calculate electronic coupling values the multistate generalized Mulliken-Hush (GMH) method developed by Cave and Newton has been employed.^{100,102} Both 2-state and 3-state variations of GMH have been used to calculate the V^2 values.

The projection of the dipole moments along the charge transfer vector (which is defined as the vector connecting the mass centers of the two molecules) is used to calculate the couplings. The adiabatic dipole moment matrix \mathbb{M}_{ad} can be diagonalized and the \mathbb{M}_{diab} matrix is obtained:

$$\mathbb{U}^T \mathbb{M}_{ad} \mathbb{U} = \mathbb{M}_{diab} \quad (7.2)$$

Notice that the diabatic dipole moments that belong to the two states of $\text{Ind}^{\bullet+}$ approximately have the same value (μ_{GS} and μ_{LE}). The third one (μ_{CT}), with a different value, belongs to the $\text{G}^{\bullet+}$ state, Table 7.1.

The eigenvector matrices \mathbb{U} and \mathbb{U}^T are obtained during the diagonalization of \mathbb{M}_{ad} . These eigenvector matrices can transform the adiabatic energy matrix into the diabatic Hamiltonian matrix (\mathbb{H}):

$$\mathbb{U}^T \mathbb{E} \mathbb{U} = \mathbb{H} \quad (7.3)$$

The off-diagonal terms of \mathbb{H} represent the electronic coupling (V) between the states.

To obtain the Hamiltonian matrix is a common procedure for both 2 and 3-state methods. The coupling calculation differs from one method to the other at this point.

7.4.1 2-state GMH method

Applying the 2-state model, electronic coupling can be expressed via the vertical excitation energy also called adiabatic splitting ($E_2 - E_1$), the difference of the adiabatic dipole moments ($\mu_2 - \mu_1$), and the transition dipole moment (μ_{12}).

$$V_{da} = \frac{(E_2 - E_1)|\mu_{12}|}{\sqrt{(\mu_2 - \mu_1)^2 + 4\mu_{12}^2}} \quad (7.4)$$

7.4.2 3-state GMH method

If we take into account that electronic coupling has a physical meaning only in intermolecular interactions; the coupling between the two states from Indole radical cation must be zeroed. In order to achieve this, a block diagonalization matrix (\mathbb{B}) is used. The diabatic Hamiltonian matrix (\mathbb{H}) is then transformed into a pseudo-diabatic Hamiltonian matrix (\mathbb{H}').

$$\mathbb{B}^T \mathbb{H} \mathbb{B} = \mathbb{H}' = \begin{pmatrix} H_{GS} & V_{GS,CT} & 0 \\ V_{CT,GS} & H_{CT} & V_{CT,LE} \\ 0 & V_{LE,CT} & H_{LE} \end{pmatrix} \quad (7.5)$$

In equation 7.5, GS is the Ground State of the system where the hole is on the indole fragment (Ind^+). LE is the Locally Excited state with also the cationic charge located on indole (Ind^+). CT is the Charge Transfer diabatic state where the hole is located on guanine (G^+). Figure 7.3 is a scheme of the employed states. In the pseudo-diabatic Hamiltonian matrix (eq. 7.5), the off diagonal terms correspond to the electronic coupling between $\text{G}^{\bullet+}$ and the two states of $\text{Ind}^{\bullet+}$ ($V_{GS,CT}$, $V_{LE,CT}$). As can be seen the $V_{GS,LE}$ coupling between the $\text{Ind}^{\bullet+}$ states is 0.

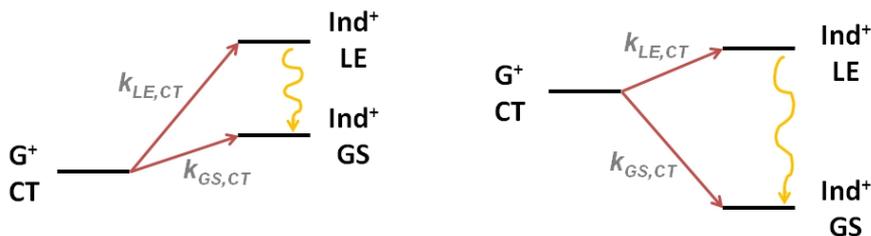


Figure 7.3: For $(\text{G-Ind})^{\bullet+}$ systems, 3 states can be involved in the charge transfer reaction. One from Guanine (donor) and two from Indole (hole acceptor).

When is the 3-state treatment necessary?

To apply a 2-state method in the calculation of electronic coupling, the two first states of the system (CT and GS) should be complementary. So as, the sum of these two first adiabatic dipole moments must be a multiple of the theoretical dipole moment (μ_{id}). This theoretical dipole moment is estimated using the point-charge model for a perfect electron transfer,

$\mu_{id}=er_{DA}$ where e is the electron charge. For each conformation, r_{DA} is approximated as the distance between the centers of mass of G and Ind. For these (G-Ind) $^{\bullet+}$ systems μ_{id} is between 6.5 and 7.8 debye (Table 7.1). Thus, π_1 , π_2 and π_3 will be much better described taking into account the first three states.^{90,101,104}

A simple diagnostic method, proposed by Cave et al.,¹⁰⁶ can be used to predict when a third state should be considered in the calculation of the electronic coupling element for a given pair of diabatic states, within the context of the generalized Mulliken-Hush approach. The diagnostic is formulated on the basis of Löwdin partitioning theory. Thus, to employ the Cave diagnostic method the transition dipole moments of the three adiabatic states considered are needed (eq. 7.5).

$$\left| \frac{\frac{\mu_{GS,LE}\mu_{LE,CT}}{\mu_{CT}-\mu_{LE}}}{\mu_{GS,CT}} \right| = |\lambda_D| \ll 1 \quad (7.6)$$

When the λ_D value (Eq. 7.6) is much less than one the 2-state method can be applied.

Taking into account a third state in the calculations of coupling also allows to calculate the HT rate between the CT from guanine and the LE state of indole. If the difference of energy between the CT and LE states is small, and they have strong coupling, their reaction rate ($k_{LE,CT}$) can have a considerable contribution to the total HT rate of the system.

$$k_{HT} = k_{GS,CT} + k_{LE,CT}$$

7.5 Results and discussion

In order to study the HT reaction between several (G-Ind) $^{\bullet+}$ systems, transition dipole moments (μ) are used to estimate the electronic coupling (V). The MS-CASPT2 level of theory with an active space of (11,12) is used to obtain both dipoles (μ) and the difference of energy between the states (ΔE).

7.5.1 Dipole moments and energies

Dipole moments are projected between the centers of mass of G and Ind, the distance between them is r_{DA} . r_{DA} takes as initial point guanine's

center of mass. From μ diabatic values (Table 7.1) the different states can be identified. The state with $\mu_{diab} \approx 0$ is the $G^{\bullet+}$ state (CT), because r_{DA} should be close to zero. The other two states which have similar values, around 7 debye, are the two $Ind^{\bullet+}$ states considered in the calculations (GS and LE).

Table 7.1: MS-CASPT2 difference of adiabatic states energy (in eV) compared to E_0 state for (G-Ind) $^{\bullet+}$ structures. Adiabatic state dipole moment (μ_{ad}), diabatic state dipole (μ_{diab}) and theoretical state dipole moment (μ_{id}) of Charge Transfer, Ground State and Locally Excited states. Dipole moments in debye.

Structure	State	ΔE_{ad}	μ_{ad}	ΔE_{diab}	μ_{diab}	μ_{id}
$\pi 1$	E_0 CT	0	4.323	0	0.698	0
	E_1 GS	0.115	4.865	-0.169	6.884	6.563
	E_2 LE	0.742	5.233	0.331	6.840	6.563
$\pi 2$	E_0 GS	0	5.917	-0.345	7.517	7.751
	E_1 LE	0.423	5.350	0.056	7.195	7.751
	E_2 CT	0.688	3.663	0	0.218	0
$\pi 3$	E_0 CT	0	5.082	0	1.696	0
	E_1 GS	0.207	6.168	-0.016	8.706	7.714
	E_2 LE	0.463	7.293	0.280	8.140	7.714
$\pi 4$	E_0 GS	0	5.037	-0.191	6.634	7.805
	E_1 CT	0.340	0.913	0	-0.707	0
	E_2 LE	0.530	7.232	0.265	7.255	7.805
$\pi 5$	E_0 GS	0	4.537	-0.146	6.892	6.842
	E_1 CT	0.449	2.297	0	-0.090	0
	E_2 LE	0.528	6.398	0.230	6.430	6.842
$\pi 6$	E_0 CT	0	1.316	0	-0.667	0
	E_1 GS	0.043	4.498	0.012	6.278	6.942
	E_2 LE	0.271	5.916	0.244	6.120	6.942

As said in the previous section, the sum of the first two adiabatic dipole moments in some cases is not close to the theoretical dipole moment (μ_{id}). $\pi 4$ and $\pi 6$ are more or less at 1 debye to accomplish it, while for the $\pi 1$, $\pi 2$, and $\pi 3$ structures the first 2 two states are clearly not complementary. This could be an evidence that a 2-state method cannot properly describe these (G-Ind) $^{\bullet+}$ systems. This dipole moment analysis is confirmed by the Cave's diagnostic method¹⁰⁶ (Table 7.2). The λ_D values of the 3 first structures are clearly not much lower than 1 ($\pi 1$ and $\pi 2$ structures have $\lambda_D > 1$), while $\pi 4$, $\pi 5$, and in lesser extent $\pi 6$ are suitable for a 2-state treatment to calculate electronic coupling.

Table 7.2: λ_D values for the Cave’s diagnostic method to predict multistate effects on the calculation of electron transfer parameters using the Generalized Mulliken-Hush method.¹⁰⁶

Structure	λ_D
$\pi 1$	1.7517
$\pi 2$	1.0779
$\pi 3$	0.3608
$\pi 4$	0.0002
$\pi 5$	0.0077
$\pi 6$	0.0434

Depending on the π -stacked conformation studied the ordination of electronic states can change (Figure 7.3 shows two possible ordination of the states). The state with $\mu_{diab} \approx 0$, the CT state, is taken as a reference to calculate the ΔE_{diab} . ΔE_{diab} is used in the 3-state model V calculations and in the HT rate calculations. Instead of diabatic energy, ΔE_{ad} is used in the calculation of V when the 2-state model is applied. The structural changes between the 6 studied conformations give significant differences between the energies of the states. For the reaction between CT-GS states, the maximum difference between these states is of -0.345 eV for the $\pi 2$ structure, while the minimum is of 0.012 eV and is obtained for $\pi 6$. If the charge transfer reaction between CT-LE states is considered, the maximum difference of energy is of 0.331 eV and belongs to $\pi 1$ structure while the minima is of 0.056 eV by $\pi 2$ (Table 7.1).

7.5.2 Electronic couplings

Analyzing electronic coupling values obtained between GS-CT states (Figure 7.4), both 3-state and 2-state methods give similar results except for $\pi 2$ structure which has the 2-state V^2 value much stronger than the obtained with the 3-state method. This difference is of 0.046 eV², so the 2-state method overestimates the coupling. However, the 2-state method does not have a clear tendency to overestimate the coupling, for example for $\pi 1$ structure the value obtained with the 3-state model is 0.0012 eV² stronger than the 2-state one.

For couplings between CT-LE states, large differences between the 3-state and 2-state V^2 are observed for the first 3 structures. As for the GS-CT case, these errors do not present a tendency. For $\pi 1$ and $\pi 3$ the 2-state model overestimates the coupling (0.041 and 0.025 eV² respectively), while for $\pi 2$ it is underestimated by 0.006 eV².

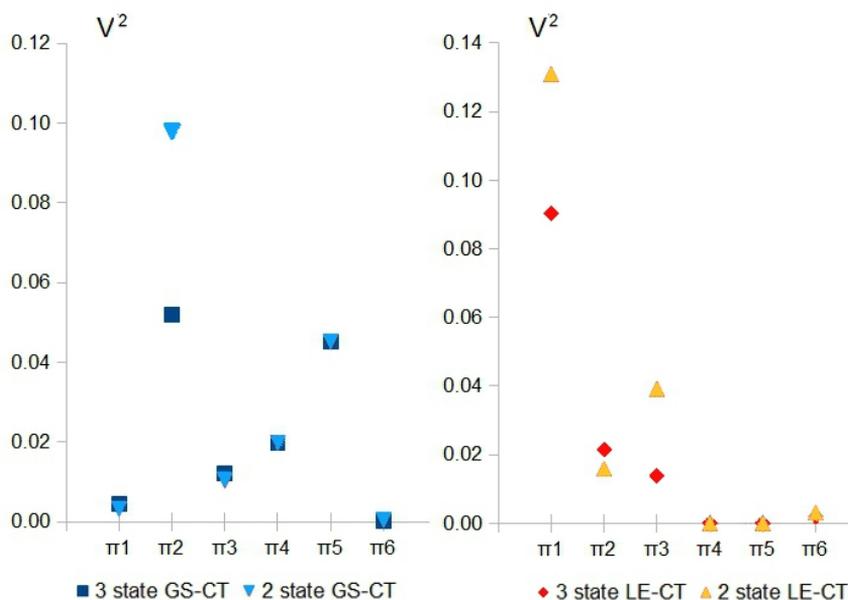


Figure 7.4: Graphical representations of electronic coupling square values (V^2) in eV^2 . 3-state and 2-state models have been employed to calculate electronic coupling, both between Ground and Charge Transfer states (GS-CT), and between the Locally Exited and the Charge Transfer state (LE-CT).

Comparing the V^2 values obtained for LE-CT and GS-CT states (Figure 7.5), strong LE-CT couplings are obtained by the first 3 structures. For π_1 and π_3 , LE-CT couplings are even stronger than GS-CT ones. This also happens for π_6 , even though their couplings are much weaker, its $V^2_{LE,CT}$ value is of 0.002 eV^2 while the $V^2_{GS,CT}$ is of 0.0003 eV^2 . These results are in agreement with the previous observations of the states involved in the calculation (dipole moment analysis) and with Cave's diagnostic method (Table 7.2). For the first 3 structures, a 2-state method to calculate electronic coupling will not describe correctly the guanine-indole interaction.

7.5.3 HT rates with 3-state method

Strong electronic couplings between LE and CT states could produce fast $k_{LE,CT}$, if the difference of energy between LE and CT states is small. This can improve the total rate of the hole transfer reaction ($k_{HT} = k_{GS,CT} + k_{LE,CT}$). Employing Marcus equation (eq. 7.1), both $k_{GS,CT}$

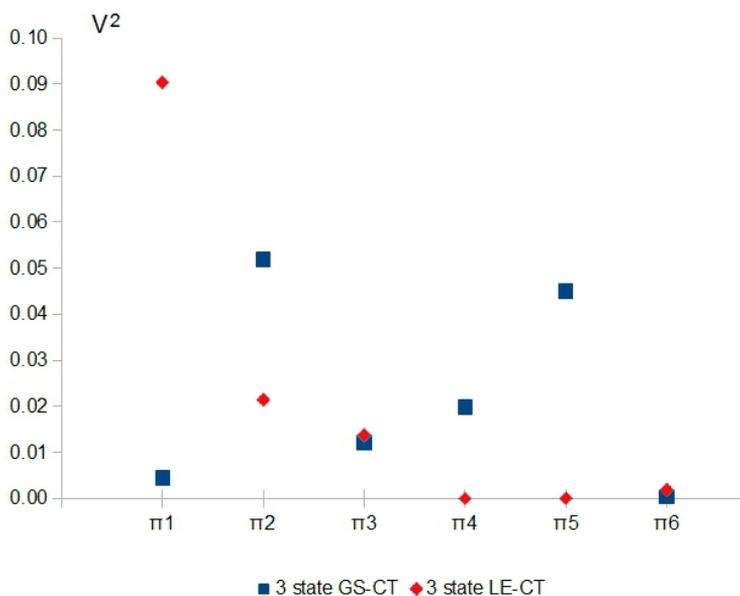


Figure 7.5: Electronic coupling square values calculated using 3-state model, for 6 different $(\text{G-Ind})^{\bullet+}$ stacked conformations. Both GS-CT and LE-CT charge transfer reactions are shown.

and $k_{LE,CT}$ are calculated. Electronic couplings values obtained with the 3-state method were used to calculate the rates (Figure 7.6). For the rate calculations, the temperature of the system is set to 298.15 K and the reorganization energy is 1 eV, as usually employed for charge transfer reactions in biomolecules.

Fast hole transfer rates are obtained for $(\text{G-Ind})^{\bullet+}$ conformations. GS-CT rates are always faster than LE-CT ones, as could be assumed (Figure 7.6). However, in some cases the rates of LE-CT are considerably fast, such as the one obtained for π_2 structure which has a $k_{LE,CT}$ of $7.02 \times 10^9 \text{ s}^{-1}$. Thus, taking into account the locally excited state of indole in the HT reaction will substantially modify the total rate of the reaction. However, for π_4 and π_5 conformations the rates obtained for the HT reaction between LE and CT states are significantly smaller than the GS-CT ones. This is not unexpected since our predictions said that a 2-state method will correctly describe its HT reactions. Thus, the LE state contribution at the total rate of the HT reaction will be negligible.

Considering the coupling between GS-CT states, the faster rate is obtained

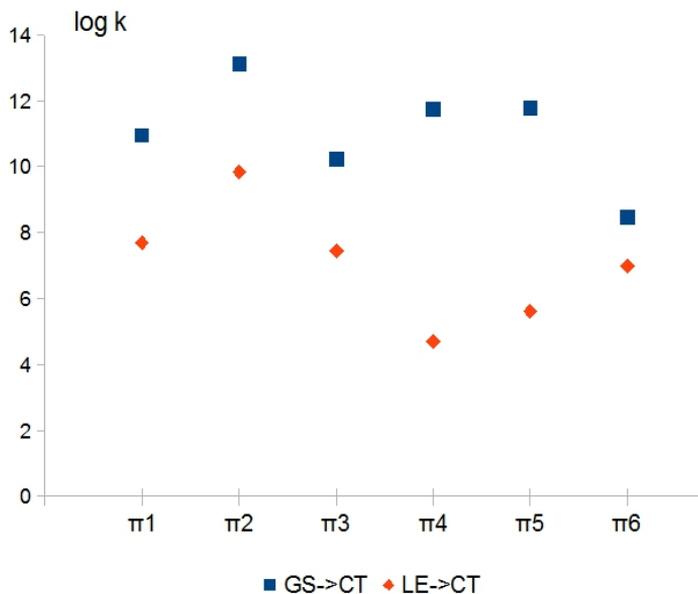


Figure 7.6: Rate constant logarithm of two HT possible reactions, $k_{GS,CT}$ and $k_{LE,CT}$. Values of electronic coupling were obtained with the 3-state model. The reorganization energy value used to calculate the rate employing the Marcus equation was of 1 eV, and the temperature was set to 298.15 K.

by $\pi2$ conformation ($1.34 \times 10^{13} \text{ s}^{-1}$), which is 2 orders of magnitude faster than the closest rate. $\pi6$ is the conformation less favorable to produce the HT reaction, with a HT rate of $2.91 \times 10^8 \text{ s}^{-1}$. $\pi6$ structure also has the HT rate for the coupling between its LE-CT states closer than the other conformations, with a difference between them of 2 orders of magnitude. Considering both $k_{GS,CT}$ and $k_{LE,CT}$ the $\pi2$ structure is the most favorable to perform the HT reaction.

7.5.4 Comparison $k_{2\text{-state}}/k_{3\text{-state}}$

The HT rates can be calculated using the values of electronic couplings obtained with the 2-state or the 3-state methods. The difference between the rates obtained with the two methods can be easily compared looking at their ratio (Table 7.3). Moreover at comparing the ratio between the rates, the effect of the approximation employed for the reorganization energy ($\lambda=1 \text{ eV}$) is canceled.

The 2-state and 3-state methods are in good agreement for HT rates

Table 7.3: Ratio between the rates of hole transfer reactions (between CT-GS or CT-LE states) calculated with the 2-state or the 3-state method.

Structure	$k_{2-state}/k_{3-state}$	$k_{2-state}/k_{3-state}$
	CT-GS	CT-LE
$\pi 1$	0.03	86.62
$\pi 2$	0.87	0.42
$\pi 3$	0.30	7.57
$\pi 4$	1.00	10.35
$\pi 5$	0.97	23.35
$\pi 6$	0.82	1.39

between CT-GS states. Looking at the ratio of rates obtained by the hole transfer reactions between LE and CT states, obtained values show that the 2-state method presents large errors. Thus, the 2-state method cannot be used in the calculations of the electronic coupling between guanine and the locally excited state of the indole radical cation. However, rate constants are not drastically affected by the use of a 2 state method. For case $\pi 2$ there is an error compensation: V is larger for 2-state model, but the 3-state model is more exothermic (See Table A.15). Overall similar values are obtained at calculating the rate with both methods.

7.6 Conclusions

The results provided in this work show that, to calculate the hole transfer parameters (electronic coupling and free energy) of guanine-indole stacked systems, a 3-state model is necessary to guarantee a proper description of the system. Large errors in the calculation of $V_{LE,CT}$ are obtained if the 2-state method is employed. This conclusion shows that the HT behavior of (G-Ind) \bullet^+ systems strongly deviates from the one of HT inside DNA, where the 2-state model can be employed.

The inclusion of another state in the calculations (the locally excited state of indole radical cation) allows to calculate a second electronic coupling between the molecules. This second coupling, called $V_{LE,CT}$, is strong enough to produce rates around 10^4 - 10^9 sec $^{-1}$. These rates for the indole excited state are significantly slower than the ones obtained for the ground state, thus $k_{LE,CT}$ can be neglected.

Chapter 8

Hole Transfer in Guanine-Indole Systems: Comparison of electronic coupling obtained with DFT and MS-CASPT2 levels of theory

In preparation

8.1 Abstract

The study of hole transfer reactions on (G-Ind) \bullet^+ system is especially challenging because of the proximity of indole radical cation ground and excited states. Several (G-Ind) \bullet^+ conformations are difficult to be accurately described only accounting for 2-state model. A set of DFT methods have been used to calculate their electronic coupling values. They were tested using as a reference data electronic coupling values obtained with 3-state GMH and MS-CASPT2 calculations.

Electronic coupling data obtained with high level CASPT2 methodology are compared with values obtained using HF and several DFT functionals (LC-wPBE, CAM-B3LYP, and wB97XD). Long-range corrected functionals and M06-2X are found to perform an accurate description of electronic coupling values for several (G-Ind) \bullet^+ conformations.

8.2 Introduction

Charge transfer reactions on Guanine-Tryptophan (G-Trp) \bullet^+ system can play a protective role against DNA oxidative damage.⁴³ In previous studies, (G-Trp) \bullet^+ system was difficult to model due to the proximity between the ground and first excited state of Indole radical cation, Ind \bullet^+ , which is a fundamental moiety of Trp directly involved in the charge transfer reaction.¹³⁰⁻¹³² A computational effort has been done and MS-CASPT2 calculations of (G-Ind) \bullet^+ charge transfer capabilities were performed (see Chapter 7), which were used as benchmarks to assess the accuracy of widely employed 2-state electronic coupling calculations. In this study, MS-CASPT2 calculations are compared with electronic couplings obtained by HF and DFT functionals: BLYP, B3LYP, PBE0, and M06-2X.¹³⁴ Also, Long range Correlated (LC) functionals have been tested (LC-wPBE, wB97X-D, and CAM-B3LYP). These LC-functionals should describe charge transfer reactions more accurately than regular functionals.

In this study, not only π -stacked structures have been considered (see A.4), but T-shaped ones (see A.4) have been also taken into account (Figure 8.1). Usually these T-shaped structures present weak couplings.

In previous work, see chapter 7, electronic coupling between the first excited state of Ind \bullet^+ (LE) and guanine charge transfer state has also been considered. This LE-CT coupling can have a significant influence on

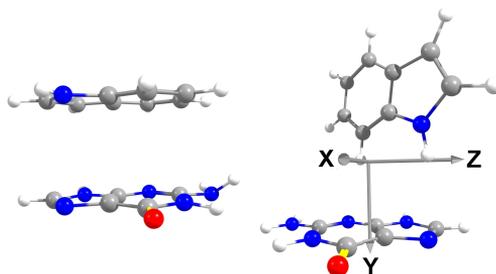


Figure 8.1: π 1 stacked conformation and T3 (T-shaped) Guanine-Indole structures.

the total rate. In the present study, this LE-CT coupling have also been calculated at DFT level of theory.

8.3 Computational details

MS-CASPT2 calculations are taken as reference, they are compared with HF and several DFT functionals results. Namely DFT functionals are BLYP, B3LYP, and M06-2X. Also LC-functionals have been employed (LC-wPBE, wB97X-D, and CAM-B3LYP).^{135–137} These LC-functionals have the computational cost of a regular functional, but they have the self-interaction error corrected. They are specially well suited to describe long-range charge transfer reactions and also non-covalent interactions. However, LC functionals have tended to be inferior to the best hybrids for properties such as thermochemistry. For electronic coupling calculations, direct method (see section 2.1.1) has been used, except for CASPT2 benchmark data which have been calculated using 2-state and 3-state GMH method (see section 2.1.4).

The basis set cc-pVTZ has been used in all HF and DFT calculations using Gaussian 03.¹³⁸

CASPT2 calculations have been performed with the ano-s basis set and the program MOLCAS 7.2.¹³³ For the calculations of $(\text{G-Ind})^{\bullet+}$, active spaces of 11 electrons in 12 orbitals (11,12) have been used. These twelve orbitals of the active space were divided equally between the two molecules to describe their π orbitals (6 orbitals for G and 6 for Ind), which are associated with charge transfer reactions.

8.3.1 LC-DFT theoretical details

For a regular hybrid functional, the exchange-correlation potential (ϵ_{xc}) decays asymptotically as $-a/r$, where a is the fraction of HF exchange. However, this decay must be an exact $-1/r$ asymptote. To correct this error a range separation can be introduced. Thus, a long-range portion of the regular functional is replaced by the HF counterpart.¹³⁹ This can be implemented by splitting the Coulomb operator into a short-range (SR) and long-range (LR) parts (equation 8.1).

$$\frac{1}{u} = \underbrace{\frac{\text{erf}(wu)}{u}}_{SR} + \underbrace{\frac{\text{erfc}(wu)}{u}}_{LR} \quad (8.1)$$

The non-Coulomb part of exchange functionals typically dies off too rapidly and becomes very inaccurate at large distances, making them unsuitable for modeling processes such as electron excitations to high orbitals. Moreover, common density functionals have problems at describing systems with a noninteger number of electrons, which is often the case in charge transfer reactions.^{140,141} Various schemes have been devised to handle such cases.

The LC-Functionals used in this work are:

LC-wPBE is remarkably accurate for a broad range of molecular properties, such as thermochemistry, barrier heights of chemical reactions, bond lengths, and most notably, description of processes involving long-range charge transfer.¹³⁵

CAM-B3LYP this functional combines the hybrid qualities of B3LYP and the long-range correction presented by Tawada et al.¹⁴² CAM-B3LYP yields atomization energies of similar quality to those from B3LYP, while also performing well for charge transfer excitations in a dipeptide model, which B3LYP underestimates enormously. The CAM-B3LYP functional comprises of 0.19 Hartree-Fock (HF) plus 0.81 Becke 1988 (B88) exchange interaction at short-range, and 0.65 HF plus 0.35 B88 at long range. The intermediate region is smoothly described through the standard error function with parameter 0.33.¹³⁶

wB97X-D includes 100% long-range exact exchange, a small fraction (about 22%) of short-range exact exchange, a modified B97 exchange density functional for short-range interaction, the B97 correlation density functional, and empirical dispersion corrections. wB97X-D provides significant improvement only for non-covalent interactions.¹³⁷

8.4 Results and discussion

Electronic coupling values (V) of 8 different (G-Ind) $^{\bullet+}$ conformations have been calculated with the direct method at HF, DFT, and CASPT2 level of theory. As DFT concerns, some LC-DFT functionals have been considered.

As it was discussed in previous chapter, when using CASPT2 calculations a 2-state or 3-state approximation can be used to obtain electronic coupling values. The values obtained with the 3-state GMH method using CASPT2 have been used as benchmark values.

The results section is divided in two parts. The first part analyzes the electronic coupling between the Charge transfer state of G $^{\bullet+}$ (CT) and the Ground State of Ind $^{\bullet+}$ (GS). The second section presents results obtained for the electronic coupling between the Charge Transfer (CT) state of G $^{\bullet+}$ and the Locally Excited State of Ind $^{\bullet+}$ (LE). All data are collected in Table 8.1 and Table 8.2 and represented in Figures 8.3 and 8.2.

8.4.1 V_{GS-CT}

Analyzing V_{GS-CT} values obtained for the interaction between the charge transfer state and the ground state (Figure 8.2), the HF method tends to overestimate V , while using DFT they are underestimated, specially BLYP, which is the worst functional at reproducing the benchmark values with an average relative error of -36.30%. M06-2X is the best performing functional (without long range correction), with an average error of -14.14%, having similar accuracy to two of the three LC-DFT methods tested (CAM-B3LYP and wB97XD). However, the LC-DFT functional LC-wPBE has been shown to be the more precise.

Comparing electronic coupling values obtained by CASPT2 calculations using the 2-state direct method with reference values (obtained with the 3-state GMH method), the 2-state values are in quite good agreement, particularly π_4 , π_5 and π_6 . However, for π_1 , π_2 and π_3 the error is significantly large. This is due to the coupling with the locally excited state of Ind radical cation. In the previous chapter (see Table 7.2), Cave diagnostic method predicted that the 3-state treatment is necessary for these three (G-Ind) $^{\bullet+}$ structures. For these structures, the HOMO and HOMO-1 orbitals of Ind are close in energy and this results in a mixing of orbitals for the G-Ind system. The adiabatic states employed in direct method calculations are not a linear combination of diabatic states. This makes the direct method less accurate.

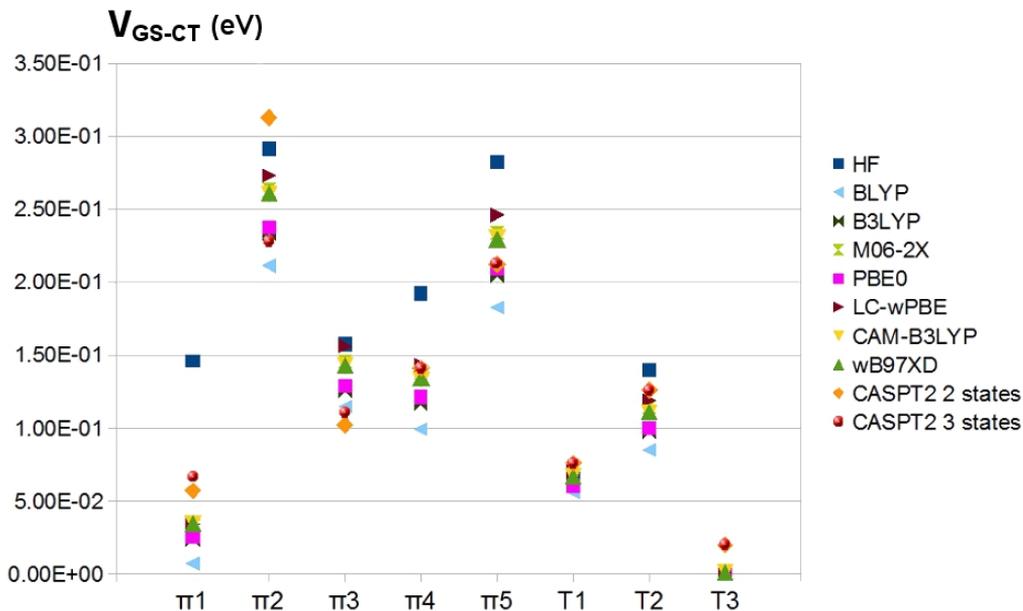


Figure 8.2: Comparison of electronic coupling values between GS-CT states V_{GS-CT} .

The electronic coupling values obtained with DFT functionals present approximately the same tendency respect the CASPT2 3-state values, depending on the conformation they can underestimate or overestimate the coupling. Also depending on the conformation functionals change their accuracy. Thus, BLYP strongly underestimate $\pi 1$ V value respect the benchmark result, like B3LYP method does in a lesser extent. For $\pi 3$ structure, it is B3LYP that is less accurate than BLYP. $\pi 2$ and $\pi 3$ are complicated conformations, thus all the functionals are overestimating their electronic coupling instead of their usual tendency to underestimate it. Only HF follows a clear tendency of overestimating electronic coupling.

Table 8.1: Electronic coupling obtained for (G-Ind)^{•+} system between ground and charge transfer states, in meV.

Structure	HF	BLYP	B3LYP	M06-2X	PBE0	LC-wPBE	CAM-B3LYP	wB97XD	CASPT2 2s	CASPT2 3s
$\pi 1$	145.70	7.14	23.50	32.65	25.93	33.85	34.33	34.25	56.95	66.58
$\pi 2$	291.50	211.40	233.10	263.70	237.60	273.00	260.40	261.20	312.90	228.04
$\pi 3$	157.60	114.20	125.30	145.20	128.80	156.30	142.70	142.80	102.02	110.43
$\pi 4$	192.20	99.01	116.50	138.20	121.20	143.00	133.30	134.20	140.73	140.69
$\pi 5$	282.60	182.20	204.70	234.20	209.40	246.20	230.80	229.40	212.38	212.37
T1	67.11	56.35	60.34	66.91	60.90	69.55	67.04	66.99	75.69	75.69
T2	139.70	85.03	96.82	114.40	100.10	119.10	110.60	111.30	125.93	125.93
T3	1.27	0.91	0.94	0.58	1.04	0.60	0.89	1.00	19.69	19.69

Table 8.2: Electronic coupling obtained for (G-Ind)^{•+} system between locally excited and charge transfer states, in meV.

Structure	HF	BLYP	B3LYP	M06-2X	PBE0	LC-wPBE	CAM-B3LYP	wB97XD	CASPT2 2s	CASPT2 3s
$\pi 1$	348.80	215.10	244.60	284.70	251.60	289.20	274.10	273.20	361.94	300.63
$\pi 2$	192.20	49.55	68.84	87.43	72.56	87.65	83.62	83.21	125.84	146.57
$\pi 3$	141.30	72.89	83.32	95.91	84.90	103.70	97.92	98.42	198.01	117.76
$\pi 4$	4.92	37.15	36.04	37.61	36.35	36.49	36.64	36.90	5.60	4.25
$\pi 5$	54.59	43.80	33.01	30.24	31.12	31.13	30.03	30.19	6.92	7.86
T1	45.19	17.20	22.03	28.31	23.32	32.99	28.05	29.30	51.95	51.95
T2	33.45	5.65	1.46	0.22	0.99	0.29	1.22	1.67	31.51	31.51
T3	12.26	7.92	8.71	9.60	8.71	9.67	9.62	9.61	19.51	19.51

8.4.2 V_{LE-CT}

Significant differences of electronic coupling values obtained with CASPT2 2 and 3-state methods are found for the coupling between Locally Exited state of $\text{Ind}^{\bullet+}$ and the Charge transfer state of $\text{G}^{\bullet+}$. However, this only happens for $\pi 1$, $\pi 2$ and $\pi 3$ structures. In previous chapter, this behavior has been analyzed. $\pi 1$, $\pi 2$ and $\pi 3$ structures require a 3-state treatment to enhance the accuracy.

In most conformations HF values are much more accurate than the ones obtained with DFT and LC-functionals. Moreover, HF is also better than the 2-state CASPT2 calculations for V_{LE-CT} values of $\pi 1$, $\pi 3$ and $\pi 4$ structures. The relative error between the direct coupling calculated with DFT methods and the reference CASPT2 3-state method is huge for $\pi 5$ structure and large for $\pi 4$ and T2 conformations (Figure 8.3). The benchmark V_{LE-CT} values of these conformations are considerably weak (Table 8.2). Thus, small deviations of LC-functionals values respect CASPT2 3-state data strongly modulates their accuracy.

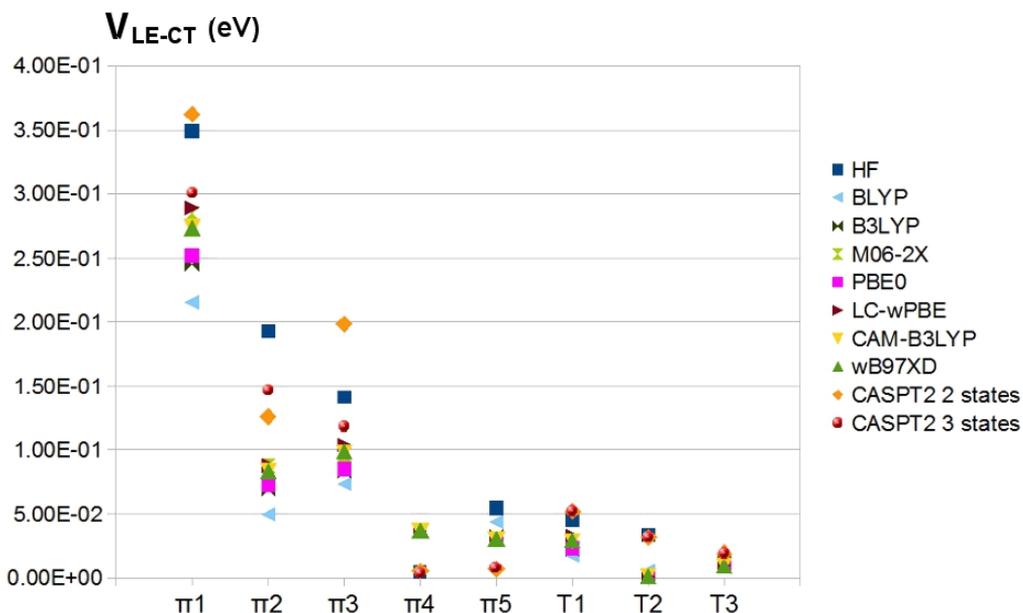


Figure 8.3: Comparison of electronic coupling values between LE-CT states, V_{LE-CT} .

8.5 Conclusions

Considering the electronic coupling between the ground state of indole and the charge transfer state of guanine, HF overestimates electronic coupling values while, in general, DFT functionals underestimate it. LC-wPBE is the most accurate functional to calculate electronic coupling. It is in good agreement with CASPT2 calculations using 2 or 3-state methods.

Similar results are found for electronic coupling values between locally excited and charge transfer states (LE-CT). HF values are remarkably accurate. Large errors are obtained for structures that do not need a 3-state treatment ($\pi 4$, $\pi 5$, and T-shaped ones).

Further analysis should be done in CT rate calculations considering also energy data.

Chapter 9

Results and discussion

The main interest of this thesis is to study charge transfer reactions within DNA-protein interactions. These reactions have been modeled by nucleobase-amino acid interactions. Therefore all possible nucleobase-aromatic amino acid complexes have been analyzed taking into account different types of conformations such as π -stacked and T-shaped ones. Their electronic coupling values (which strongly modulates charge transfer rate) and hole transfer rates have been calculated (Chapter 4).

The next stage of the thesis evaluates the effect of conformational fluctuations on Guanine-Tryptophan and Adenine-Tryptophan electronic coupling values (Chapters 5 and 6). Then the accuracy of the widely used 2-state method to calculate electronic coupling values has been assessed in comparison with a 3-state method. For this study the calculations have been performed at MS-CASPT2 level of theory (Chapter 7).

Finally, using as benchmark electronic coupling values obtained with 3-state calculations at MS-CASPT2 level of theory, several computational methods (HF, DFT) have been tested in order to analyze their accuracy (Chapter 8).

9.1 Electron transfer from aromatic amino acids to guanine and adenine radical cations in π -stacked and T-shaped complexes

If DNA is oxidized, the cationic charge created can migrate through the stack of nucleobases till it forms a mutagenic lesion. This cationic charge can be transferred to an interacting protein or peptide by a charge transfer reaction. The redox properties of the aromatic amino acids make them the most favorable ones to stabilize the cationic charge by transferring an electron to DNA.

Using Marcus charge transfer theory, together with direct 2-state method to calculate electronic coupling, the electron transfer capabilities of Guanine-X and Adenine-X systems (where X=Phe, His, Tyr, and Trp) have been analyzed. The study has been carried out at DFT level of theory, but using structures optimized at MP2 level by Wetmore *et al.*¹⁴³ Studied G-X and A-X conformations were not exclusively π -stacked interactions, T-shaped ones have also been considered (Figure 9.1).

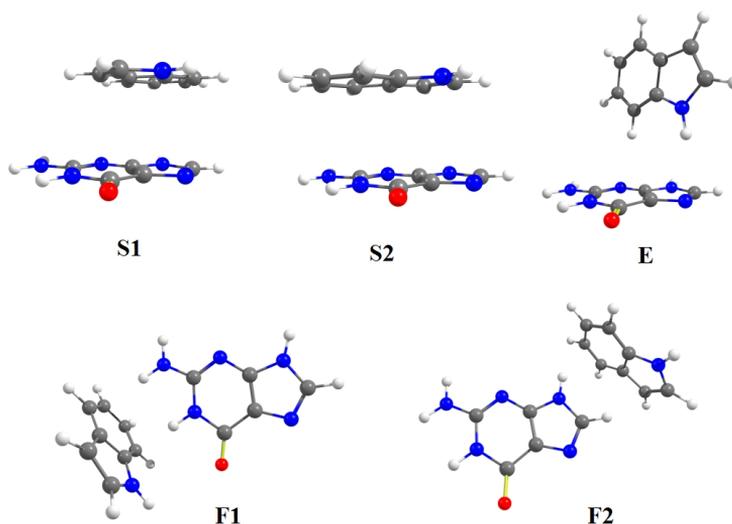


Figure 9.1: Structure of the guanine-tryptophan complexes: stacked conformations S1 and S2, and T-shaped conformations E (edge) and F1 and F2 (face).

9.1.1 G-X and A-X distribution of charge

Considering the ground state of G-X and A-X radical cations, the distribution of charge gives information of where the cationic charge is localized. In Figure 9.2, the difference of charge between the molecules (ΔQ) is displayed.

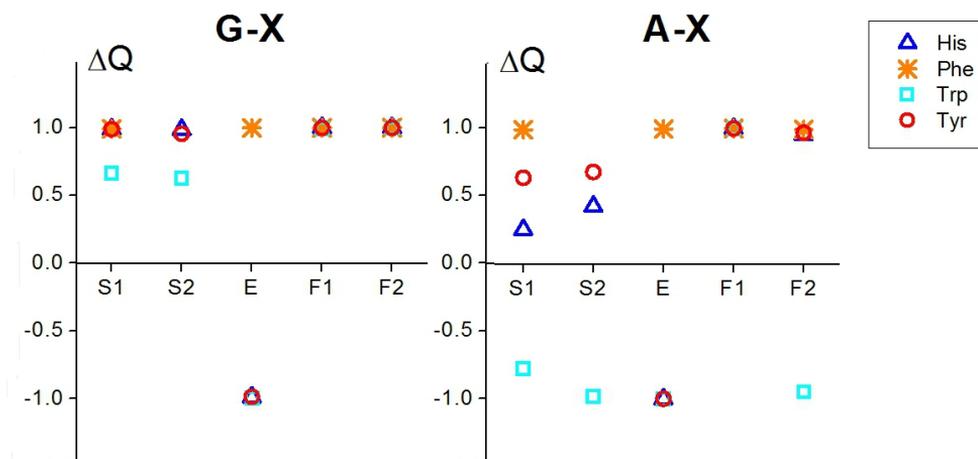


Figure 9.2: Charge difference between the nucleobase and residue. $\Delta Q = Q(N) - Q(X)$ in radical cations $(G-X)^{\bullet+}$ and $(A-X)^{\bullet+}$, where $X = \text{Trp, Tyr, His, and Phe}$. $\Delta Q = 1$ charge on the NB. $\Delta Q = -1$ charge on the amino acid

Considering A-X structures, high variability is observed in the distribution of charges. All A-Trp interactions have the charge localized on tryptophan. For stacked dimers of A-Tyr and A-His, the charge is delocalized. Phenylalanine has the charge always located on adenine so this amino acid does not seem as favorable as the others to conduct charge transfer reactions with this nucleobase.

The variations on the charge localization depending on the conformation (specially for the A-X systems) suggest that a charge transfer reaction can be triggered by structural changes.

ΔQ is related with other charge transfer parameters (which are calculated in the following sections) such as the electronic coupling (V) and the driving force (ΔE) (see eq. 9.1).

$$\Delta Q = \frac{\Delta E}{\sqrt{\Delta E^2 + 4V^2}} \quad (9.1)$$

For G-X systems, the cationic charge is mainly confined on guanine because their ΔE values usually are positive. Only on edge conformations the charge is located in the amino acid, except for phenylalanine interaction. For tryptophan amino acid, some delocalization has been found in π -stacked conformations. This is because their values of electronic coupling and driving force are similar.

9.1.2 G-X and A-X electronic couplings and free energy

Electronic coupling values have shown to be extremely sensitive to the mutual arrangement of monomers (Figure 9.3). Not only by comparing between π -stacked and T-shaped conformations, even between stack conformations itself there could be significant variations of electronic coupling value. One example of this sensibility can be seen on the G-His interaction.

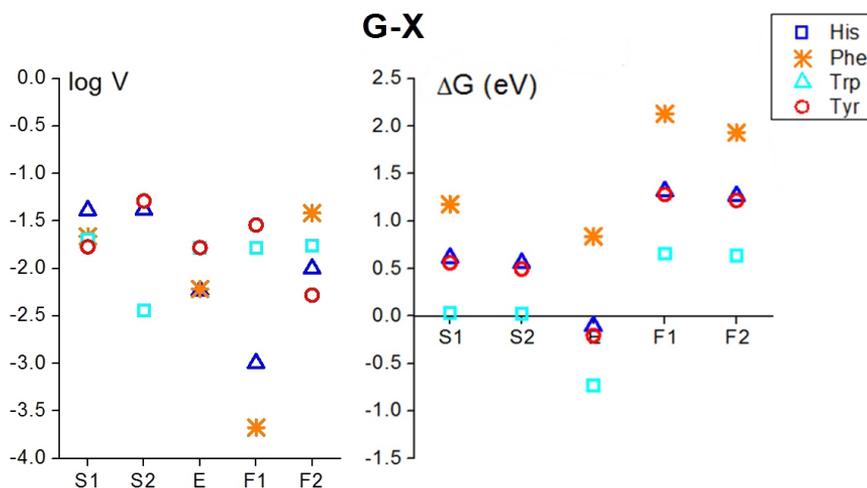


Figure 9.3: Dependence of the electronic coupling and free energy on the dimer conformation for G-X systems (X=Trp, Tyr, His and Phe).

Moreover, the dependence of electronic coupling on the configurations can not be predicted without using computational calculations. For some compounds, T-shaped interactions can have stronger couplings than π -stacked ones, this also depends on the aromatic amino acid which interacts with the nucleobase. No clear conformational preference to enhance the electronic coupling has been observed.

Unexpected strong couplings have been found for T-shaped interactions.

It is usually assumed that π -stacking between the molecules is required to produce a charge transfer reaction, however, the obtained results refute this assumption.

In most G-X structures, ΔG (which is approximated by ΔE , see section 1.3) is positive (Figure 9.3), and therefore, the ET process is unlikely. Negative ΔG values are found in the E conformation of G-X, where X=Trp, Tyr and His. As the ionization energy of A is by 0.4 eV larger than that of G, the A^+ state can be reduced more easily. Independent of the conformation of G-Phe and A-Phe, the ET driving force is calculated to be positive in these complexes. As expected, Trp is the best reducing agent among the aromatic amino acids. Tyr and His have very similar ionization energies.

9.1.3 G-X and A-X hole transfer rates

Hole transfer rates do not only depends on the electronic coupling, there are also other parameters such as the driving force and reorganization energy. The reorganization energy used to be in a range of 0.5 to 1.5 eV in biological systems. For these calculations, a reorganization energy of 1 eV has been considered. The driving force of the systems also depends on the conformation, however, all the systems present the same tendency in their variations.

Some of obtained rates are faster than 10^8 sec^{-1} . The guanine radical cation deprotonation produce guanine radicals $[G(-H)^\bullet]$ which are a first step to the formation of permanent damage on DNA. Experimentally the rate of the deprotonation reaction has been calculated to be of 10^7 sec^{-1} .¹⁴⁴ Thus, it seems feasible to extract a cationic charge from DNA before the guanine radical formation. Both His, Trp, and Tyr aromatic amino acids are suitable to obtain fast electron transfer rates at interacting with guanine or adenine.

9.2 Conformational Dependence of the Electronic Coupling in Guanine-Tryptophan and Adenine-Tryptophan Complexes: A DFT Study

As has been observed in section 9.1, electronic coupling values of Guanine or Adenine interactions with aromatic amino acids suffer drastic variations between the studied conformations. In order to determine the nature of these variations, a systematic conformational study has been carried out for G-Trp and A-Trp interactions. For example, analyzing (G-Trp) $^{\bullet+}$ charge

distribution (which depends on electronic coupling) when shift and slide translations are applied (Figure 9.4), large variations are observed for few angstroms displacements. Thus, charge transfer reactions on $(\text{G-Trp})^{\bullet+}$ can be triggered by conformational changes. In the following sections, the conformational dependence of parameters that modify the charge transfer rate (electronic coupling and driving force) will be evaluated for $(\text{G-Trp})^{\bullet+}$ and $(\text{A-Trp})^{\bullet+}$ systems.

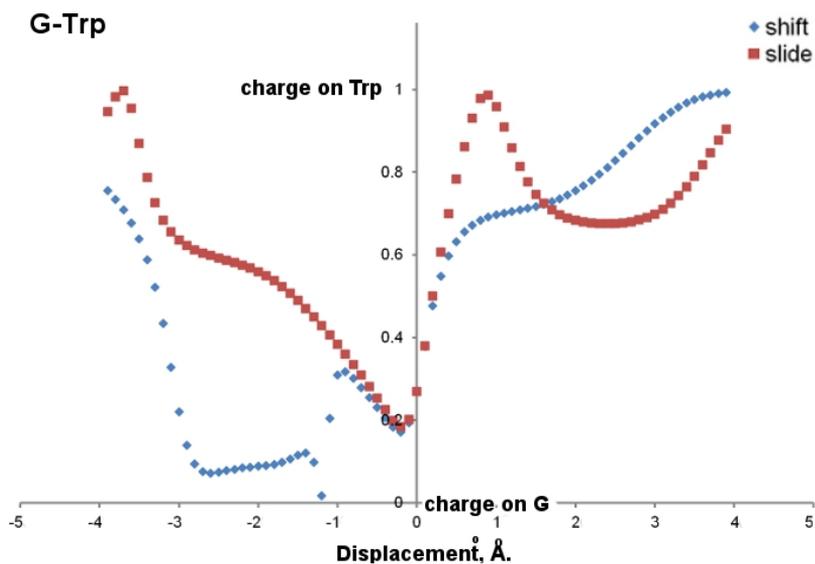


Figure 9.4: Dependence of Charge distribution for G-Trp interactions obtained after applying displacements on shift and slide parameters. A value of 1 means that the charge is exclusively located on Trp, while for a value of 0 it is located on Guanine.

Adenine is not as stable as hole acceptor as can be guanine. However, for long bridges between guanine nucleobases in DNA, adenine can also act as a stepping stone of multistep charge transfer process. In Chapter 4 electronic coupling values calculated for a set of A-Trp interactions have been shown to be stronger than the structurally related G-Trp ones.

Using DFT calculations and the 2-state direct method to estimate electronic coupling values, five of the six parameters that define the mutual arrangement of the subunits (see section 2.5.1) have been explored. The sixth parameter called the rise, which is the distance between the centers of mass of the molecules, have not been analyzed because it is known that

electronic coupling decrease exponentially when the distance between the molecules grows.

The obtained results can be divided between π -stacked conformations (formed by variations on shift, slide, or twist parameters) and T-shaped ones (obtained with movements on roll or tilt parameters, or also displacements applied to a T-Shaped structure).

9.2.1 Electronic couplings of π -stacked conformations

Shift and Slide

Strong electronic coupling values have been found for π -stacked conformations (Figure 9.5). Only few structures present V values below $V < 0.01$ eV, what we have considered the limit of a weak interaction. The difference of energy between the perfect stacked conformation and structures obtained after the displacements, on shift and slide transitional parameters, do not have a direct influence on the electronic coupling between the molecules.

The V results of stacked interactions present a counterintuitive behavior.

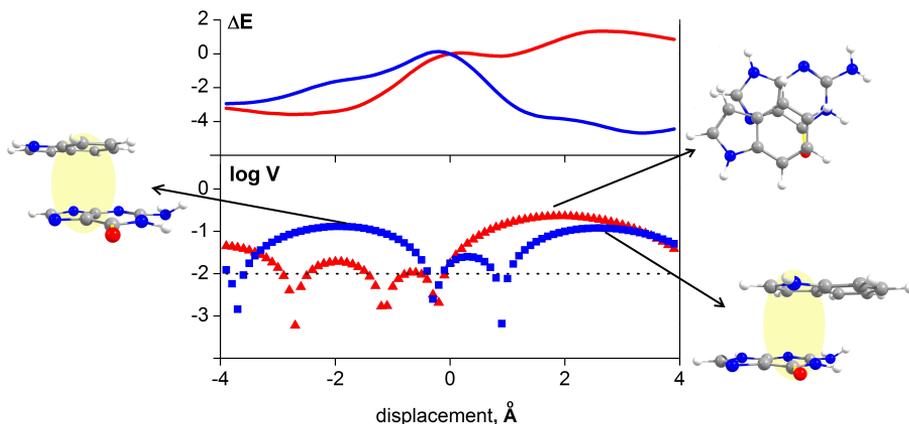


Figure 9.5: Dependence of the conformational energy (ΔE in kcal/mol) and the electronic coupling ($\log V$) on Trp displacements on shift (red) and slide (blue) parameters. The value of $V=0.01$ eV ($\log V=-0.2$) is shown by the dashed line represents the boundary between values considered strong or weak.

Strongest values were not obtained when the overlap between the aromatic rings of the molecules is maximized. A partial stacking of only one ring is preferable for slide movement. A symmetric behavior for slide electronic

coupling values is observed. The first maximum (Figure 9.5) is obtained when the stacking is mostly with the 6-member Trp ring with guanine or adenine while the second maximum is found when the 5-member ring of Trp is stacked with the nucleobase.

For G-Trp movements on shift parameter, the strongest electronic coupling value is found when there is almost no stacking between the molecules (Figure 9.5). For A-Trp, the maximum V shift value is also obtained for a conformation where the direct overlap between the rings is considerably small.

Twist

Strong values were obtained both for G-Trp and A-Trp systems, but for angles close to 0 and 180 degrees, couplings were extremely weak. It seems that the position of the C-C bond adjacent to the 6 member Trp ring enhance the electronic coupling when it is over polarized bonds of guanine, then strongest electronic coupling values were obtained.

Comparing the electronic coupling results obtained both for G-Trp and A-Trp interactions the importance of polarized bonds to maximize the electronic coupling is more evident (Figure 9.6). G-Trp and A-Trp have a V maximum for a rotation of 110 degrees (structures 1 and 2 on Figure 9.6). These two conformations are structurally highly similar, the C-C bond adjacent to the 6 member ring of Trp is over a polarized bond. It seems that the C-N bond of Adenine enhances better than the C=O of guanine the electronic coupling of the system. Guanine has another polarized bond (C-N) that produce a V maximum when it is over the same C-C bond of tryptophan (structure 3 on Figure 9.6). However, the coupling obtained for this conformation is not as strong as the obtained over the C=O bond. So it is not so clear that a C-N bond enhances more the coupling, other conformational factors have to be taken into account.

9.2.2 Electronic couplings of T-shaped G-Trp and A-Trp conformations

Roll and Tilt

Analyzing electronic coupling values of T-shaped conformations (Figure 9.7) obtained after roll or tilt rotations, it is observed that they are in general much weaker than π -stacked ones. However, for angles between 45 and 130 degrees that belong to conformations more T-shaped than stacked, V values become stronger.

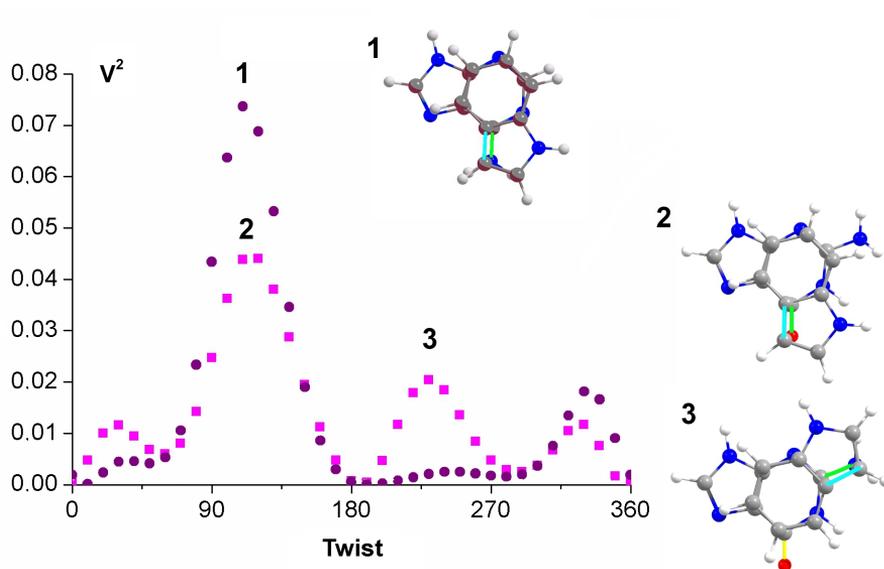


Figure 9.6: Comparison of electronic coupling values (eV^2) obtained for A-Trp (purple) and G-Trp (pink) interactions at applying twist rotations (in degrees).

Notice that V tilt values between 50 and 120 degrees are not represented (although they are remarkably strong) because the destabilization of these structures is more than 5 kcal/mol.

Translations along the X, Z, and XZ axis have been applied to G-Trp and A-Trp T-shaped structures optimized by Wetmore *et al.*¹⁴³ at MP2 level of theory. Mostly the obtained conformations (G-Trp ones shown in Figure 9.8) do not present strong electronic coupling values but some of them are quite strong to ensure a HT reaction. High variability of electronic coupling is observed for these T-shaped conformations.

9.3 Hole Transfer in Guanine-Indole Systems: A Multi-Configurational Study

The Indole (Ind) moiety of tryptophan has its locally excited (LE) state very close in energy to its ground state (GS). A minimum difference between these diabatic states was found to be of 0.232 eV. Thus, this second state of Ind can participate in the HT reaction. The widely used 2-state method

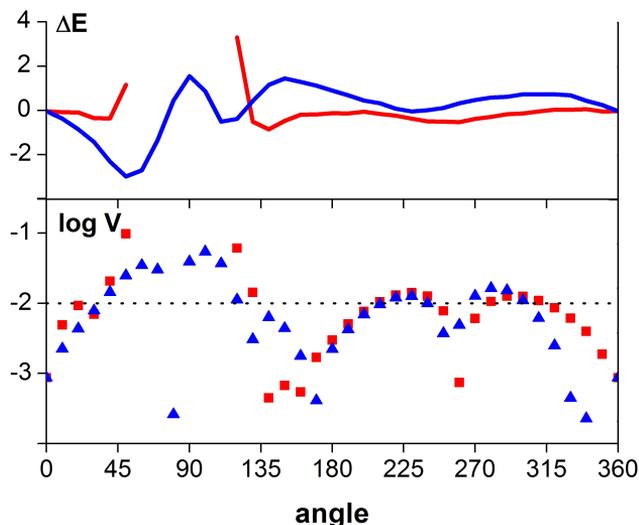


Figure 9.7: Dependence of the conformational energy (ΔE in kcal/mol) and the electronic coupling on rotations (in degrees) of Trp in the stacked complex: the data for tilt and roll are shown in red and blue, respectively. The value of $V=0.01$ eV ($\log V=-0.2$) is shown by the dashed line represents the boundary between values considered strong or weak.

to calculate electronic coupling may be of limited use for systems involving Ind radical cation.

MS-CASPT2 calculations with an active space of (11,12) have been performed in several π -stacked (G-Ind) $^{\bullet+}$ interactions in order to apply a 3-state method to calculate the hole transfer parameters and compare their results with the ones obtained using a 2-state method. Within the 3-state method, the second state of Ind $^{\bullet+}$ and their coupling with the Charge Transfer (CT) state of guanine (V_{LE-CT}) can be calculated (see section 2.1.4). Cave's diagnostic method¹⁰⁶ predicts when a 2-state method is insufficient to calculate the electronic coupling of the system within the use of Generalized Mulliken-Hush method (see section 2.1.4 for more details). The results obtained for this diagnostic method are presented in Table9.1. Taking into account the Cave's diagnostic method prediction, the set of π -stacked structures can be divided in two groups:

- $\pi 1$, $\pi 2$, and $\pi 3$** These structures present values of λ_D that are not much smaller than 1. A 2-state model can not properly describe the electronic coupling of these structures.

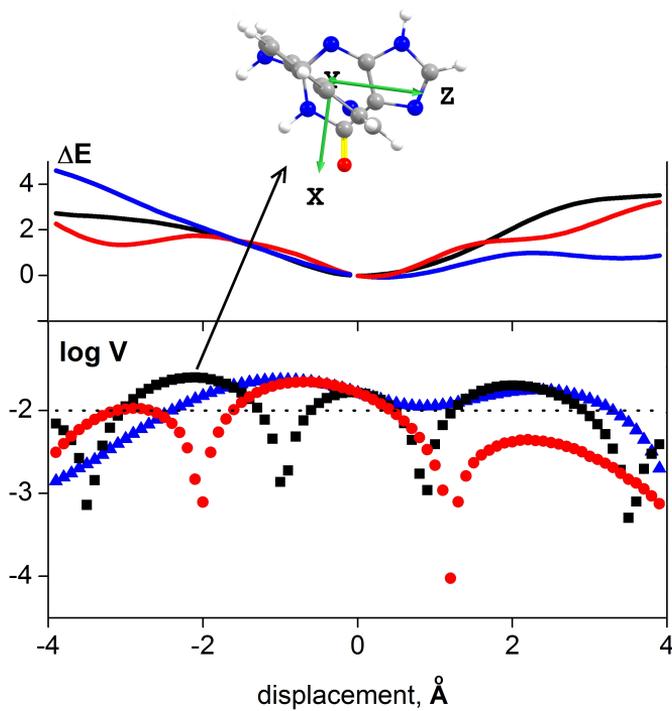


Figure 9.8: Dependence of the conformational energy (ΔE in kcal/mol) and electronic coupling on displacement of Trp in the T-shaped complexes: the data for X, Z, and XZ are shown in blue, black, and red, respectively. The value of $V=0.01$ eV ($\log V=-0.2$) is shown by the dashed line represents the boundary between values considered strong or weak.

Table 9.1: λ_D values for the Cave's diagnostic method to predict multistate effects on the calculation of electron transfer parameters using the Generalized Mulliken-Hush method¹⁰⁶

Structure	λ_D
$\pi 1$	1.7517
$\pi 2$	1.0779
$\pi 3$	0.3608
$\pi 4$	0.0002
$\pi 5$	0.0077
$\pi 6$	0.0434

$\pi 4$, $\pi 5$, and $\pi 6$ The values of λ_D are much smaller than 1 so a 2-state model can correctly describe these conformations.

Similar conclusions can be found by analyzing the adiabatic dipole moment.

9.3.1 Electronic couplings of G-Ind systems

Comparing the electronic coupling values obtained for the interaction between GS-CT states, both 3-state and 2-state methods are in good agreement except for $\pi 2$ structure which electronic coupling are largely overestimated by the 2-state method (see Figure 9.9).

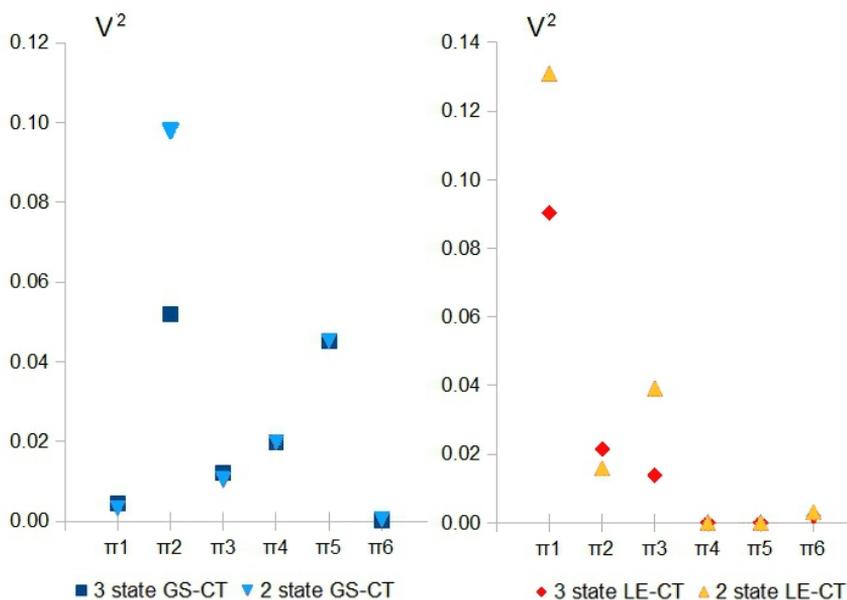


Figure 9.9: Graphical representations of electronic coupling square values (V^2) in eV^2 . The 3-state and the 2-state models to calculate V have been employed, both between GS-CT and LE-CT states.

For the first three structures ($\pi 1$, $\pi 2$, and $\pi 3$), strong electronic coupling values were obtained between the LE and CT states. This was expected with the Cave's diagnostic prediction. However, the obtained electronic coupling values using the 2-state method for this LE-CT coupling do not even have a qualitative meaning comparing them with the 3-state results. The 2-state method can overestimate or underestimate the coupling depending on the structure. If considering the coupling of LE-CT states and calculating it with the 2-state method would give reasonable coupling values, the use of less computationally expensive levels of theory to obtain the transition dipole moments used in the calculations of electronic coupling would be possible. However, a multi-configurational method is clearly

required to describe charge transfer reactions in systems involving indole radical cation.

9.3.2 Hole transfer rates using the 3-state method

Hole transfer rates have been calculated using Marcus equation 1.4. The total rate of the HT reaction will be the sum of two possible reactions, depending on which states are interacting:

$$k_{HT} = k_{GS-CT} + k_{LE-CT}.$$

As could be expected, the rates obtained for the reaction with the ground state of Ind are faster than the ones obtained with its locally excited state (Figure 9.10). However, for the HT between LE and CT several fast rates were obtained. For example, $\pi 2$ conformation has a k_{LE-CT} value of $7.018 \times 10^9 \text{ s}^{-1}$. These fast rates can have a significant contribution to the total rate of the reaction.

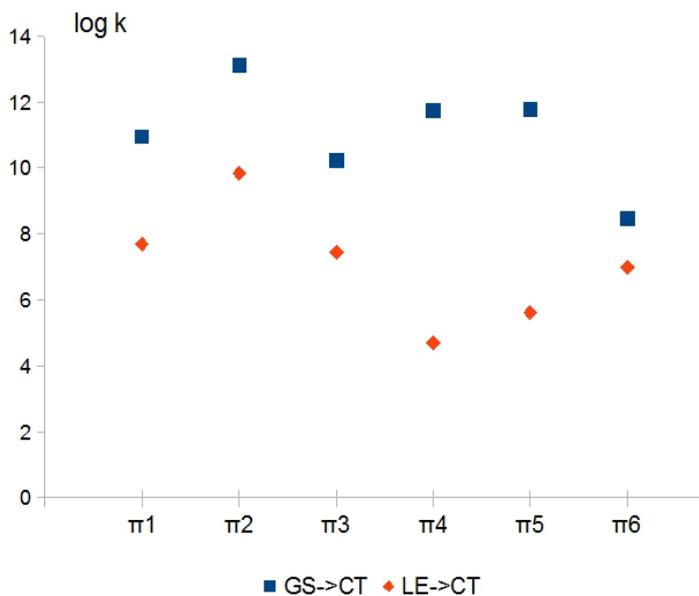


Figure 9.10: Rate constant logarithm of two HT possible reactions, k_{GS-CT} and $k_{LE,CT}$. V values were obtained with 3-state GMH. The λ used was of 1 eV, and a temperature of 298.15 K.

Although $\pi 6$ structure has been included on the structures where the 2-state model can be applied, notice that this conformation presents the minimum difference between the rates obtained for GS-CT and LE-CT interactions (Figure 9.10).

9.4 Hole Transfer in Guanine-Indole Systems: Comparison of electronic coupling obtained with DFT and MS-CASPT2 levels of theory

Electronic coupling values (V) of 8 different (G-Ind)^{•+} conformations have been calculated with the direct method and using HF, DFT and LC-DFT functionals, and CASPT2 methodologies. The accuracy of these calculations have been analyzed using benchmark V values obtained with a 3-state GMH method using the MS-CASPT2 level of theory.

9.4.1 V_{GS-CT}

Analyzing V_{GS-CT} values obtained for the interaction between guanine charge transfer state and indole ground state, the HF method trends to overestimate V , while using DFT V values are underestimated, specially using BLYP, which is the worst functional at reproducing the benchmark values (see Table 8.1). M06-2X is the best performing functional (without long range correction) with similar accuracy than CAM-B3LYP and wB97XD LC-functionals. LC-wPBE has been the most accurate DFT method with an average relative error in comparison to the benchmark data of -10.10%.

Comparing electronic coupling values obtained by CASPT2 calculations using the 2-state direct method with the reference values obtained with the 3-state GMH method, the 2-state values are extremely accurate. However, for $\pi 1$, $\pi 2$ and $\pi 3$ the error is significant. This is due to the coupling with the locally excited state of Ind radical cation. Thus, the adiabatic states employed in the electronic coupling calculations are not linear combinations of diabatic states. This was shown by dipole moment analysis.

9.4.2 V_{LE-CT}

Electronic couplings calculated for LE-CT states have larger errors than couplings between GS-CT states. Comparing V values obtained with DFT and direct method and reference CASPT2 3-state GMH calculation, the accuracy is extremely poor for $\pi 4$, $\pi 5$ and T2 conformations. The benchmark V_{LE-CT} values of these conformations are considerably weak. Thus, small deviations of LC-functionals values for these structures strongly modulates its error.

As expected, V_{LE-CT} values obtained with the 2-state CASPT2 method (direct method) substantially differ from the ones obtained with the 3-state

GMH.

Surprisingly in most conformations, HF values are much more accurate than the ones obtained with DFT and LC-functionals. Moreover, HF is also better than the 2-state CASTP2 calculations for $\pi 1$, $\pi 3$ and $\pi 4$ structures.

Electronic coupling values obtained with DFT methods does not present a clear tendency (like the HF one to overestimate electronic coupling) compared to benchmark data. For some structures quite accurate results can be obtained but small conformational changes lead to large errors. When dealing with (G-Ind)^{•+}, a 3-state GMH method and a CASPT2 level of theory seems the only reliable method to calculate its electronic coupling.

Not only electronic coupling modulates the charge transfer rate, but also driving force. Further work will be done in this direction.

Chapter 10

Conclusions

The main conclusions of this thesis are divided in four sections.

The electron transfer reaction between guanine or adenine and the four aromatic amino acids (Chapter 4) have been studied. Both π -stacked and T-shaped conformations have been considered in this study. Electron transfer parameters, driving force and electronic coupling, were calculated at DFT level of theory. We obtained the following conclusions:

First: Strong electronic coupling values were found for π -stacked structures. They will ensure effective electron transfer from aromatic amino acids to $G^{\bullet+}$ or $A^{\bullet+}$. Surprisingly we obtained quite strong couplings also for T-shaped conformations, where the two molecules are perpendicular. Thus, stacking between NB and the amino acid is not required for a HT reaction.

Second: The driving force is strongly dependent on the NB-aaa conformation. The most favorable values for this parameter were obtained for edge structures, a type of T-shaped conformations.

Third: The excess charge and the spin density of NB-X structures are usually confined to a single subunit. Depending on the conformation it could be on the nucleobase or on the aromatic amino acid. Thus, structural changes could trigger the charge transfer reaction between the sites.

Fourth: The obtained rates show that the ET reaction from Trp to $G^{\bullet+}$ can be faster than the deprotonation of $G^{\bullet+}$. This deprotonated guanine form is precursor of highly mutagenic species. So, if the cationic charge can be extracted from DNA before permanent damage formation it may play crucial role protecting the genetic material.

Conclusions for the study of the conformational dependence of electronic coupling both for G-Trp (Chapter 5) and A-Trp (Chapter 6) interactions are:

Fifth: G-Trp and A-Trp systems are particularly sensitive to conformational changes. Small fluctuations of less than 1 Å can change the HT rate by several orders of magnitude. Thus, to evaluate the HT reactions in a biological environment thermal fluctuations should be considered and molecular dynamics simulations should be performed, calculating electronic coupling as an average of possible conformations. To analyze the HT capabilities of a NB amino acid pair the use of a single conformation, such a crystallographic structure, can lead to erroneous conclusions.

Sixth: For studied G-Trp and A-Trp pairs total eclipsed conformations of aromatic rings produce weak coupling. Strongest couplings were obtained when only one of tryptophan's aromatic rings has a π - π interaction with the nucleobase. This behavior is substantially different than the one observed by symmetric structures of NB pairs (G-G, A-A), where a perfect overlap between the aromatic rings of the molecules leads to the strongest electronic coupling.

Seventh: Although usually T-shaped structures have weaker couplings than π -stacked ones, some conformational regions with the molecules almost perpendicular present values large enough to ensure HT reactions. T-shaped NB-aaa structures could also play an important role on biological charge transfer reactions.

For the multi-configurational CASPT2 study of hole transfer reactions in (G-Ind) \bullet^+ systems (Chapter 7) applying both the general 2-state method and the 3-state one, the following conclusions were obtained:

Eighth: To calculate the hole transfer parameters (electronic coupling and free energy) of guanine-indole stacked systems, a 3-state model is necessary to guarantee a proper description of the system. The HT behavior of (G-Ind) \bullet^+ systems strongly deviates from the one of HT inside DNA, where the 2-state model can be employed.

Ninth: The inclusion of another state in the calculations (the locally excited state of indole radical cation) allows to calculate a second electronic coupling between the molecules, $V_{LE,CT}$. It is strong enough to produce rates around 10^4 - 10^9 sec $^{-1}$. However, these rates are significantly slower than the ones obtained for the ground state, thus $k_{LE,CT}$ can be neglected.

In the comparison between different levels of theory to describe (G-Ind)^{•+} charge transfer reactions (Chapter 8), electronic coupling values obtained have been analyzed. The conclusions can be summarized in:

Tenth: Considering the electronic coupling between the ground state of indole and the charge transfer state of guanine, HF overestimates electronic coupling values while DFT methods underestimate it. LC-wPBE is the most accurate functional to calculate electronic coupling. It is in good agreement with CASPT2 calculations using 2 or 3-state methods.

Eleventh: Similar results are found for electronic coupling values between locally excited and charge transfer states. HF values are remarkably accurate. Larger errors are obtained for structures which do not need a 3-state treatment (π 4, π 5, and T-shaped ones).

List of Acronyms

a Acceptor

aaa Aromatic Amino Acid

A Adenine

CASSCF Complete Active Space Self-consistent Field

CASPT2 The Møller-Plesset perturbation theory using a CASSCF function as a reference wave function.

CI Configuration Interaction

CSF Set of Configuration Functions

CT Charge Transfer

d Donor

DFT Density Functional Theory

EET Excess Electron Transfer

ET Electron Transfer

FCM Fragment Charge Method

G Guanine

GMH Generalized Mulliken-Hush

GC Gradient-Corrected

GGA Generalized Gradient Approximation

GS Ground State

HF Hartree-Fock

His	Histidine
HOMO	Highest Occupied Molecular Orbital
HT	Hole Transfer
Ind	Indole
LC	Long range Correlated
LDA	Local Density Approximation
LE	Locally Excited
LSDA	Local Spin Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
MCSCF	Multi-Configurational Self Consistent Field
MO	Molecular Orbital
MP	Møller-Plesset
MS	Multi-State
NB	Nucleobase
Phe	Phenylalanine
RASSCF	Restricted Active Space Self-consistent Field
RHF	Restricted Hartree-Fock
Tyr	Tyrosine
Trp	Tryptophan
VdW	Van der Waals

Appendix A

Appendix

A.1 Supporting Material of chapter 4, "Electron transfer from aromatic amino acids to guanine and adenine radical cations in π -stacked and T-shaped complexes"

Table A.1: Notation of dimer configurations in comparison to the used by Wetmore et al.

Current study	Wetmore et al.
S1	stacked
S2	stacked
E	edge
F1	face_A for G face_4 for A
F2	face_5 for G-His G-Phe face_E for G-Trp G-Tyr face_8 for A

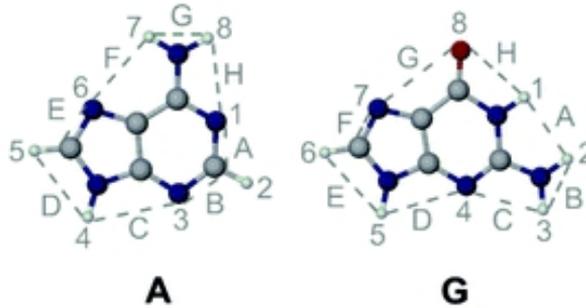


Figure A.1: Nomenclature used for Guanine and Adenine faces.

Table A.2: Comparison of the ET driving force ΔE and electronic coupling V calculated with different DFT functionals and the Hartree-Fock method for S1 [G-Trp] complex.

Functional	ΔE (eV)	V (eV)
B3LYP	0.046	0.020
BHandHLYP	0.055	0.025
PBE0	0.054	0.020
M052X	0.046	0.025
M06	0.053	0.020
M06-2X	0.058	0.026
HF	0.175	0.032

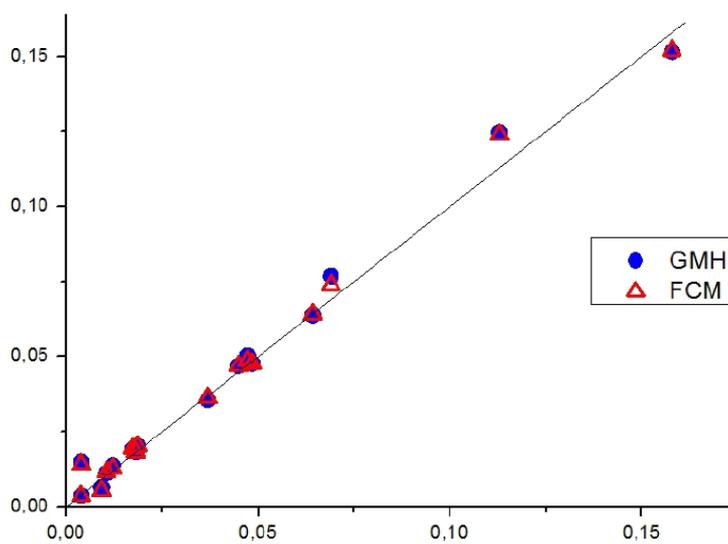


Figure A.2: Comparison of electronic couplings for A-X complexes calculated using the GMH and FCM schemes and the direct method (X-axis).

Table A.3: Computational results for complexes **G-X**. All energies in eV, electronic coupling V in meV, Charges in au. $E(G)$ is the energy in the complex and $E^0(G)$ is the isolated energy.

Complex	X	E(A)	E(A)- E^0 (A)	E(X)	E(X)- E^0 (X)	ΔE	V (meV)	Q(A)	Q(X)	K(sec $^{-1}$)
A_∞	HIS	5.514	0.000	6.092	0.000	0.578	–	–	–	–
S1	HIS	5.493	-0.021	6.108	0.016	0.615	40.930	0.994	0.006	2.70E+02
S2	HIS	5.390	-0.124	5.950	5.726	0.560	41.420	0.993	0.007	1.51E+03
E	HIS	5.910	0.401	5.805	5.581	-0.105	5.849	0.005	0.995	2.38E+08
F1	HIS	5.324	-0.185	6.643	0.550	1.319	1.004	1.000	0.000	3.21E-13
F2	HIS	5.346	-0.165	6.613	0.520	1.267	9.943	1.000	0.000	3.22E-10
∞	PHE	5.514	0.000	6.701	0.000	1.187	–	–	–	–
S1	PHE	5.370	-0.144	6.548	-0.153	1.178	21.660	0.996	0.004	7.18E-08
E	PHE	5.611	0.102	6.451	-0.250	0.840	6.057	0.999	0.001	3.07E-03
F1	PHE	5.276	-0.233	7.409	0.708	2.133	0.211	0.999	0.001	2.55E-33
F2	PHE	5.315	-0.196	7.250	0.549	1.935	38.270	1.000	0.000	9.98E-24
∞	TRP	5.514	0.000	5.384	0.000	-0.130	–	–	–	–
S1	TRP	5.312	-0.202	5.343	-0.041	0.031	20.080	0.832	0.168	2.19E+08
S2	TRP	5.304	-0.210	5.330	-0.054	0.026	3.588	0.813	0.187	7.73E+06
E	TRP	5.851	0.342	5.118	-0.266	-0.733	16.430	0.001	0.999	2.27E+12
F1	TRP	5.238	-0.273	5.896	0.512	0.658	16.470	0.998	0.002	1.11E+01
F2	TRP	5.202	-0.307	5.841	0.457	0.639	17.480	0.998	0.002	2.30E+01
∞	TYR	5.514	0.000	5.971	0.000	0.457	–	–	–	–
S1	TYR	5.397	-0.117	5.959	-0.012	0.562	16.840	0.994	0.006	2.35E+02
S2	TYR	5.373	-0.141	5.870	-0.101	0.497	51.260	0.977	0.023	1.51E+04
E	TYR	5.821	0.312	5.616	-0.355	-0.205	16.550	0.007	0.993	9.84E+09
F1	TYR	5.307	-0.204	6.590	0.619	1.283	28.580	0.999	0.001	1.31E-09
F2	TYR	5.281	-0.230	6.500	0.529	1.219	5.242	1.000	0.000	7.28E-10

Table A.4: Computational results for complexes **A-X**. All energies in eV, electronic coupling V in meV, Charges in au. $E(A)$ is the energy in the complex and $E^0(A)$ is the isolated energy.

Complex	X	E(A)	E(A)- $E^0(A)$	E(X)	E(X)- $E^0(X)$	ΔE	$V(\text{meV})$	Q(A)	Q(X)	K(sec $^{-1}$)
∞	HIS	5.876	0.000	6.092	0.000	0.216	–	–	–	–
S1	HIS	5.841	-0.035	5.976	-0.116	0.135	157.800	0.627	0.373	1.51E+09
S2	HIS	5.836	-0.040	5.934	-0.158	0.098	44.950	0.713	0.287	2.74E+08
E	HIS	6.293	0.417	5.819	-0.273	-0.474	10.470	0.001	0.999	1.25E+11
F1	HIS	5.696	-0.180	6.520	0.428	0.824	8.796	1.000	0.000	1.15E-02
F2	HIS	5.789	-0.087	6.262	0.170	0.473	68.950	0.977	0.023	5.45E+04
∞	PHE	5.876	0.000	6.701	0.000	0.825	–	–	–	–
S1	PHE	5.739	-0.137	6.550	-0.151	0.811	3.891	0.995	0.005	3.55E-03
E	PHE	5.982	0.106	6.512	-0.189	0.530	17.400	0.998	0.002	6.56E+02
F1	PHE	5.666	-0.210	7.236	0.535	1.570	18.290	1.000	0.000	7.02E-16
F2	PHE	5.708	-0.168	6.895	0.194	1.187	70.160	0.996	0.004	5.14E-07
∞	TRP	5.876	0.000	5.384	0.000	-0.492	–	–	–	–
S1	TRP	5.657	-0.219	5.357	-0.027	-0.300	112.900	0.113	0.887	1.82E+12
S2	TRP	5.720	-0.156	5.336	-0.048	-0.384	9.181	0.012	0.988	3.53E+10
E	TRP	6.221	0.345	5.179	-0.205	-1.042	18.500	0.001	0.999	5.65E+12
F1	TRP	5.630	-0.246	5.761	0.377	0.131	3.891	0.999	0.001	1.00E+06
F2	TRP	5.733	-0.143	5.515	0.131	-0.218	36.860	0.028	0.972	5.96E+10
∞	TYR	5.876	0.000	5.971	0.000	0.095	–	–	–	–
S1	TYR	5.721	-0.155	5.851	-0.120	0.130	64.260	0.818	0.182	2.80E+08
S2	TYR	5.734	-0.142	5.855	-0.116	0.121	48.360	0.839	0.161	1.93E+08
E	TYR	6.214	0.338	5.755	-0.216	-0.459	18.110	0.002	0.998	3.20E+11
F1	TYR	5.698	-0.179	6.412	0.441	0.714	12.150	1.000	0.000	9.62E-01
F2	TYR	5.747	-0.129	6.139	0.168	0.392	47.210	0.984	0.016	2.44E+05

A.2 Supporting Material of chapter 5, "Conformational Dependence of the Electronic Coupling in Guanine-Tryptophan Complexes: A DFT Study"

Table A.5: Atomic coordinates of G-Ind stacked reference structure in Angstroms.

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	-1.8653	2.2620	-3.1909
C	-1.0145	3.3122	-2.9354
C	-1.0583	1.0873	-3.3088
H	-1.2273	4.3624	-2.7814
N	0.2832	2.8464	-2.8888
C	0.2932	1.4878	-3.1133
C	-1.3308	-0.2712	-3.5602
H	1.1007	3.4168	-2.7192
C	1.3629	0.5843	-3.1612
C	-0.2749	-1.1703	-3.6087
H	-2.3539	-0.6091	-3.7140
H	2.3906	0.9085	-3.0093
C	1.0583	-0.7473	-3.4112
H	-0.4730	-2.2218	-3.8021
H	1.8618	-1.4788	-3.4559
H	-2.9418	2.3282	-3.2826

Table A.6: Atomic coordinates of G-Ind T-shaped reference structure in Angstroms.

N	-1.7225	2.4593	1.1919
C	-0.8788	2.4593	2.2783
N	0.3997	2.4593	1.9371
C	0.3720	2.4593	0.5590
C	1.4628	2.4593	-0.3811
O	2.6763	2.4593	-0.2103
N	0.9352	2.4593	-1.7114
C	-0.3906	2.4593	-2.0750
N	-0.6634	2.4593	-3.4110
N	-1.3847	2.4593	-1.2161
C	-0.9339	2.4593	0.0712
H	-1.2537	2.4593	3.2937
H	1.6613	2.4593	-2.4233
H	0.0527	2.4593	-4.1185
H	-1.6329	2.4593	-3.6884
H	-2.7357	2.4593	1.1943
H	1.7091	-1.8164	3.6220
H	0.7713	-0.1366	2.0042
H	-0.3066	-5.0875	0.1448
H	-0.6504	-0.1037	-0.4484
H	-1.4502	-4.1143	-1.8282
H	-1.6212	-1.6634	-2.1231
H	1.1577	-4.2899	2.6707
C	0.8929	-3.3450	2.2139
C	1.1850	-2.0991	2.7179
C	0.1826	-3.1460	0.9886
N	0.6903	-1.1347	1.8644
C	0.0702	-1.7410	0.7946
C	-0.3760	-4.0079	0.0250
C	-0.5740	-1.1813	-0.3165
C	-1.0151	-3.4589	-1.0775
C	-1.1133	-2.0598	-1.2470

A.3 Supporting Material of chapter 6, "Conformational dependence of the electronic coupling for hole transfer between adenine and tryptophan"

Table A.7: Atomic coordinates of A-Ind stacked reference structure in Angstroms.

N	0.3216	-2.6767	0.0000
C	1.6909	-2.7590	0.0000
H	2.2042	-3.7126	0.0000
N	2.2923	-1.5770	0.0000
C	1.2373	-0.6861	0.0000
C	1.1946	0.7233	0.0000
N	2.3216	1.4716	0.0000
H	2.2434	2.4778	0.0000
H	3.2272	1.0268	0.0000
N	0.0000	1.3367	0.0000
C	-1.1130	0.5684	0.0000
H	-2.0519	1.1181	0.0000
N	-1.2166	-0.7663	0.0000
C	0.0000	-1.3367	0.0000
H	-0.3414	-3.4429	0.0020
C	0.3654	-2.7841	3.3800
C	1.7398	-2.8383	3.3800
C	0.0000	-1.4015	3.3800
H	2.4055	-3.6920	3.3800
N	2.2465	-1.5551	3.3800
C	1.2086	-0.6502	3.3800
C	-1.2261	-0.7088	3.3800
H	3.2293	-1.3174	3.3830
C	1.2268	0.7507	3.3800
C	-1.2137	0.6787	3.3800
H	-2.1677	-1.2547	3.3800
H	2.1609	1.3091	3.3800
C	0.0000	1.4015	3.3800
H	-2.1542	1.2245	3.3800
H	-0.0270	2.4886	3.3800
H	-0.3011	-3.6370	3.3800

Table A.8: Atomic coordinates of A-Ind T-shaped reference structure in Angstroms.

C	0.9579	2.5295	-0.2287
N	1.7608	2.5295	-1.3483
C	-0.3436	2.5295	-0.7390
N	-0.3602	2.5295	-2.1197
C	0.9269	2.5295	-2.4377
C	-1.3812	2.5295	0.2156
N	-1.0641	2.5295	1.5199
C	0.2450	2.5295	1.8589
N	1.3223	2.5295	1.0644
N	-2.6841	2.5295	-0.1476
H	1.3129	2.5295	-3.4445
H	2.7686	2.5295	-1.3502
H	0.4424	2.5295	2.9243
H	-2.9354	2.5295	-1.1192
H	-3.3893	2.5295	0.5668
H	-1.7799	-1.7460	3.2067
H	-0.8477	-0.0663	1.5856
H	0.2236	-5.0172	-0.2775
H	0.5654	-0.0334	-0.8720
H	1.3604	-4.0441	-2.2544
H	1.5303	-1.5932	-2.5500
H	-1.2318	-4.2196	2.2535
C	-0.9686	-3.2747	1.7958
C	-1.2590	-2.0288	2.3008
C	-0.2626	-3.0758	0.5680
N	-0.7672	-1.0644	1.4455
C	-0.1508	-1.6707	0.3736
C	0.2926	-3.9376	-0.3975
C	0.4894	-1.1111	-0.7398
C	0.9278	-3.3887	-1.5022
C	1.0255	-1.9896	-1.6721

A.4 Supporting Material of chapter 7, "Hole Transfer in Guanine-Indole Systems: A Multi-Configurational Study"

Table A.9: Atomic coordinates of π_1 structure in Angstroms (Total stack).

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	0.3654	-2.7841	3.3800
C	1.7398	-2.8383	3.3800
C	0.0000	-1.4015	3.3800
H	2.4055	-3.6920	3.3800
N	2.2465	-1.5551	3.3800
C	1.2086	-0.6502	3.3800
C	-1.2261	-0.7088	3.3800
H	3.2293	-1.3174	3.3830
C	1.2268	0.7507	3.3800
C	-1.2137	0.6787	3.3800
H	-2.1677	-1.2547	3.3800
H	2.1609	1.3091	3.3800
C	0.0000	1.4015	3.3800
H	-2.1542	1.2245	3.3800
H	-0.0270	2.4886	3.3800
H	-0.3011	-3.6370	3.3800

Table A.10: Atomic coordinates of π_2 structure in Angstroms (Shift 2 Å).

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	2.3654	-2.7841	3.3800
C	3.7398	-2.8383	3.3800
C	2.0000	-1.4015	3.3800
H	4.4055	-3.6920	3.3800
N	4.2465	-1.5551	3.3800
C	3.2086	-0.6502	3.3800
C	0.7739	-0.7088	3.3800
H	5.2293	-1.3174	3.3830
C	3.2268	0.7507	3.3800
C	0.7863	0.6787	3.3800
H	-0.1677	-1.2547	3.3800
H	4.1609	1.3091	3.3800
C	2.0000	1.4015	3.3800
H	-0.1542	1.2245	3.3800
H	1.9730	2.4886	3.3800
H	1.6989	-3.6370	3.3800

Table A.11: Atomic coordinates of $\pi 3$ structure in Angstroms (Slide -1.7Å).

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	0.3654	-4.4841	3.3800
C	1.7398	-4.5383	3.3800
C	0.0000	-3.1015	3.3800
H	2.4055	-5.3920	3.3800
N	2.2465	-3.2551	3.3800
C	1.2086	-2.3502	3.3800
C	-1.2261	-2.4088	3.3800
H	3.2293	-3.0174	3.3830
C	1.2268	-0.9493	3.3800
C	-1.2137	-1.0213	3.3800
H	-2.1677	-2.9547	3.3800
H	2.1609	-0.3909	3.3800
C	0.0000	-0.2985	3.3800
H	-2.1542	-0.4755	3.3800
H	-0.0270	0.7886	3.3800
H	-0.3011	-5.3370	3.3800

Table A.12: Atomic coordinates of $\pi 4$ structure in Angstroms (Slide 2.7Å).

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	0.3654	-4.4841	3.3800
C	1.7398	-4.5383	3.3800
C	0.0000	-3.1015	3.3800
H	2.4055	-5.3920	3.3800
N	2.2465	-3.2551	3.3800
C	1.2086	-2.3502	3.3800
C	-1.2261	-2.4088	3.3800
H	3.2293	-3.0174	3.3830
C	1.2268	-0.9493	3.3800
C	-1.2137	-1.0213	3.3800
H	-2.1677	-2.9547	3.3800
H	2.1609	-0.3909	3.3800
C	0.0000	-0.2985	3.3800
H	-2.1542	-0.4755	3.3800
H	-0.0270	0.7886	3.3800
H	-0.3011	-5.3370	3.3800

Table A.13: Atomic coordinates of $\pi 5$ structure in Angstroms (Twist 110°).

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	2.4912	1.2956	3.3800
C	2.0721	2.6056	3.3800
C	1.3170	0.4793	3.3800
H	2.6466	3.5232	3.3800
N	0.6930	2.6429	3.3800
C	0.1976	1.3581	3.3800
C	1.0854	-0.9097	3.3800
H	0.1335	3.4851	3.3830
C	-1.1250	0.8961	3.3800
C	-0.2227	-1.3726	3.3800
H	1.9204	-1.6078	3.3800
H	-1.9692	1.5828	3.3800
C	-1.3170	-0.4793	3.3800
H	-0.4139	-2.4431	3.3800
H	-2.3293	-0.8765	3.3800
H	3.5206	0.9610	3.3800

Table A.14: Atomic coordinates of $\pi 6$ structure in Angstroms (G-Ind) $^{\bullet+}$ structure where Ind is interacting with the other guanine side.

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	-1.8653	2.2620	-3.1909
C	-1.0145	3.3122	-2.9354
C	-1.0583	1.0873	-3.3088
H	-1.2273	4.3624	-2.7814
N	0.2832	2.8464	-2.8888
C	0.2932	1.4878	-3.1133
C	-1.3308	-0.2712	-3.5602
H	1.1007	3.4168	-2.7192
C	1.3629	0.5843	-3.1612
C	-0.2749	-1.1703	-3.6087
H	-2.3539	-0.6091	-3.7140
H	2.3906	0.9085	-3.0093
C	1.0583	-0.7473	-3.4112
H	-0.4730	-2.2218	-3.8021
H	1.8618	-1.4788	-3.4559
H	-2.9418	2.3282	-3.2826

Table A.15: Electronic coupling, free energy and HT reaction rate of (G-Ind) \bullet^+ conformations. Results belongs both to the use of the 3-state (3s) and the 2-state (2s) GMH methods.

	3s GS-CT	3s GS-CT	3s GS-CT	3s LE-CT	3s LE-CT	3s LE-CT
	V (eV)	ΔG (eV)	k (s $^{-1}$)	V (eV)	ΔG (eV)	k (s $^{-1}$)
$\pi 1$	0.06658	-0.16897	8.99E+10	0.30063	0.33105	4.95E+07
$\pi 2$	0.22804	-0.34473	1.34E+13	0.14657	0.05586	7.02E+09
$\pi 3$	0.11043	-0.01620	1.66E+10	0.11776	0.27989	2.78E+07
$\pi 4$	0.14069	-0.19076	5.68E+11	0.00425	0.26470	5.28E+04
$\pi 5$	0.21237	-0.14629	6.30E+11	0.00786	0.22968	4.23E+05
$\pi 6$	0.01923	0.01225	2.91E+08	0.04477	0.24393	9.73E+06
	2s GS-CT	2s GS-CT	2s GS-CT	2s LE-CT	2s LE-CT	2s LE-CT
	V (eV)	ΔG (eV)	k (s $^{-1}$)	V (eV)	ΔG (eV)	k (s $^{-1}$)
$\pi 1$	0.05695	0.01359	2.48E+09	0.36194	0.16247	4.29E+09
$\pi 2$	0.31290	-0.28671	1.16E+13	0.12584	0.08308	2.94E+09
$\pi 3$	0.10202	0.03713	4.98E+09	0.19801	0.23971	2.11E+08
$\pi 4$	0.14073	-0.19047	5.66E+11	0.00560	0.18992	5.47E+05
$\pi 5$	0.21238	-0.14468	6.14E+11	0.00692	0.07806	9.87E+06
$\pi 6$	0.01890	0.02029	2.39E+08	0.05529	0.24768	1.36E+07

A.5 Supporting Material of chapter 8, "Hole Transfer in Guanine-Indole Systems: Comparison of electronic coupling obtained with DFT and MS-CASPT2 levels of theory"

Table A.16: Atomic coordinates of T1 structure in Angstroms (Roll 100°).

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	-0.0635	-2.7841	5.0201
C	-0.3021	-2.8383	3.6666
C	0.0000	-1.4015	5.3800
H	-0.4177	-3.6920	3.0110
N	-0.3901	-1.5551	3.1676
C	-0.2099	-0.6502	4.1898
C	0.2129	-0.7088	6.5875
H	-0.5578	-1.3174	2.1992
C	-0.2130	0.7507	4.1718
C	0.2108	0.6787	6.5753
H	0.3764	-1.2547	7.5148
H	-0.3752	1.3091	3.2519
C	0.0000	1.4015	5.3800
H	0.3741	1.2245	7.5015
H	0.0047	2.4886	5.4066
H	0.0523	-3.6370	5.6765

Table A.17: Atomic coordinates of T2 structure in Angstroms (Tilt 50°).

N	0.2668	-2.6360	-0.0040
C	1.6343	-2.7736	0.0020
H	2.1120	-3.7449	-0.0010
N	2.2691	-1.6118	0.0160
C	1.2530	-0.6819	0.0180
C	1.3268	0.7583	0.0090
O	2.2909	1.5139	-0.0030
N	0.0000	1.2919	0.0000
H	-0.0286	2.3041	-0.0980
C	-1.1756	0.5833	0.0060
N	-2.3377	1.3354	-0.0740
H	-2.3545	2.1680	0.5060
H	-3.1558	0.7530	0.0730
N	-1.2429	-0.7256	0.0150
C	0.0000	-1.2919	0.0000
H	-0.4299	-3.3719	-0.0180
C	0.3654	-1.7896	3.2473
C	1.7398	-1.8244	3.2057
C	0.0000	-0.9009	4.3064
H	2.4055	-2.3732	2.5518
N	2.2465	-0.9996	4.1887
C	1.2086	-0.4179	4.8819
C	-1.2261	-0.4556	4.8370
H	3.2293	-0.8491	4.3727
C	1.2268	0.4825	5.9551
C	-1.2137	0.4363	5.8999
H	-2.1677	-0.8065	4.4188
H	2.1609	0.8415	6.3828
C	0.0000	0.9009	6.4536
H	-2.1542	0.7871	6.3180
H	-0.0270	1.5996	7.2864
H	-0.3011	-2.3378	2.5939

Table A.18: Atomic coordinates of T3 structure in Angstroms.

N	-1.7225	2.4593	1.1919
C	-0.8788	2.4593	2.2783
N	0.3997	2.4593	1.9371
C	0.3720	2.4593	0.5590
C	1.4628	2.4593	-0.3811
O	2.6763	2.4593	-0.2103
N	0.9352	2.4593	-1.7114
C	-0.3906	2.4593	-2.0750
N	-0.6634	2.4593	-3.4110
N	-1.3847	2.4593	-1.2161
C	-0.9339	2.4593	0.0712
H	-1.2537	2.4593	3.2937
H	1.6613	2.4593	-2.4233
H	0.0527	2.4593	-4.1185
H	-1.6329	2.4593	-3.6884
H	-2.7357	2.4593	1.1943
H	1.7091	-1.8164	3.6220
H	0.7713	-0.1366	2.0042
H	-0.3066	-5.0875	0.1448
H	-0.6504	-0.1037	-0.4484
H	-1.4502	-4.1143	-1.8282
H	-1.6212	-1.6634	-2.1231
H	1.1577	-4.2899	2.6707
C	0.8929	-3.3450	2.2139
C	1.1850	-2.0991	2.7179
C	0.1826	-3.1460	0.9886
N	0.6903	-1.1347	1.8644
C	0.0702	-1.7410	0.7946
C	-0.3760	-4.0079	0.0250
C	-0.5740	-1.1813	-0.3165
C	-1.0151	-3.4589	-1.0775
C	-1.1133	-2.0598	-1.2470

Bibliography

- 1 Watson, J. D.; Crick, F. H. C. *Nature* **1952**, *171*, 737–738.
- 2 Eley, D. D.; Spivey, D. I. *Trans. Farad. Soc.* **1962**, *58*, 411–415.
- 3 Murphy, C. J.; Arkin, M. R.; Jenkins, Y.; Ghatlia, N. D.; Bossmann, S. H.; Turro, N. J.; Barton, J. K. *Science* **1993**, *262*, 1025–1029.
- 4 Arkin, M. R.; Stemp, E. D. A.; Holmlin, R. E.; Barton, J. K.; Hormann, A.; Olson, E. J. C.; Barbara, P. F. *Science* **1996**, *273*, 475–480.
- 5 Hall, D. B.; Holmlin, R. E.; Barton, J. K. *Nature* **1996**, *382*, 731.
- 6 Schuster, G. B. *Long-Range Charge Transfer in DNA: Topics in Current Chemistry*; Springer: Berlin, 2004; Vol. 236.
- 7 Porath, D.; Cuniberti, G.; Di Felice, R. *Long-Range Charge Transfer in DNA II* **2004**, 183–228.
- 8 Ariga, K.; Ji, Q.; Hill, J. In *Modern Techniques for Nano- and Microreactors/-reactions*; Caruso, F., Ed.; Advances in Polymer Science; Springer, 2010; Vol. 229; pp 51–87.
- 9 Giese, B. *Annu. Rev. Biochem.* **2002**, *71*, 51–70.
- 10 Olmon, E. D.; Sontz, P. A.; Blanco-Rodríguez, A. M.; Towrie, M.; Clark, I. P.; Vlček, A.; Barton, J. K. *J. Am. Chem. Soc.* **2011**, *133*, 13718–13730.
- 11 von Feilitzsch, T.; Tuma, J.; Neubauer, H.; Verdier, L.; Haselsberger, R.; Feick, R.; Gurzadyan, G.; Voityuk, A. A.; Griesinger, C.; Michel-Beyerle, M. E. *J. Phys. Chem. B* **2008**, *112*, 973–989.
- 12 Boon, E. M.; Barton, J. K. *Curr. Opin. Struc. Biol.* **2002**, *12*, 320–329.

- 13 Meggers, E.; Michel-Beyerle, M. E.; Giese, B. *J. Am. Chem. Soc.* **1998**, *120*, 12950–12955.
- 14 Bixon, M.; Giese, B.; Wessely, S.; Langenbacher, T.; Michel-Beyerle, M. E.; Jortner, J. *PNAS* **1999**, *96*, 11713.
- 15 Giese, B.; Amaudrut, J.; Köhler, A. K.; Spormann, M.; Wessely, S. *Nature* **2001**, *412*, 318–320.
- 16 Voityuk, A. A.; Siritwong, K.; Rösch, N. *Angew. Chem. Int. Ed.* **2004**, *43*, 624–627.
- 17 Giese, B.; Amaudrut, J.; Kohler, A. K.; Spormann, M.; Wessely, S. *Nature* **2001**, *412*, 318–320.
- 18 Bixon, M.; Jortner, J. *J. Phys. Chem. A* **2001**, *105*, 10322–10328.
- 19 Kubař, T.; Elstner, M. *J. Phys. Chem. B* **2008**, *112*, 8788–8798.
- 20 Wallrapp, F.; Voityuk, A. A.; Guallar, V. *J. Chem. Theory Comput.* **2009**, *5*, 3312–3320.
- 21 Manoj, P.; Min, C.-K.; Aravindakumar, C. T.; Joo, T. *Chem. Phys.* **2008**, *352*, 333–338.
- 22 Rink, G.; Kong, Y.; Koslowski, T. *Chem. Phys.* **2006**, *327*, 98–104.
- 23 Lewis, F. D.; Liu, X. Y.; Miller, S. E.; Haye, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2002**, *124*, 11280–11281.
- 24 Giese, B.; Carl, B.; Carl, T.; Carell, T.; Behrens, C.; Hennecke, U.; Schiemann, O.; Feresin, E. *Angew. Chem. Int. Ed.* **2004**, *43*, 1848–1851.
- 25 Ito, R. S. E., T. *Angew. Chem. Int. Ed.* **2004**, *43*, 1839–1842.
- 26 Voityuk, A. A.; Davis, W. B. *J. Phys. Chem. B* **2007**, *111*, 2976–2985.
- 27 Shinde, S. S.; Maroz, A.; Hay, M. P.; Anderson, R. F. *J. Am. Chem. Soc.* **2009**, *131*, 5203–5207.
- 28 Saito, I.; Takayama, M.; Sugiyama, H.; Nakatani, K.; Tsuchida, A.; Yamamoto, M. *J. Am. Chem. Soc.* **1995**, *117*, 6406–6407.
- 29 Roca-Sanjuán, D.; Rubio, M.; Merchán, M.; Serrano-Andrés, L.
- 30 Shih, C.; Museth, A. K.; Abrahamsson, M.; Blanco-Rodriguez, A. M.; Di Bilio, A. J.; Sudhamsu, J.; Crane, B. R.; Ronayne, K. L.; Towrie, M.; Vlček Jr, A. *Science* **2008**, *320*, 1760–1762.

- 31 Wittekindt, C.; Schwarz, M.; Friedrich, T.; Koslowski, T. *J. Am. Chem. Soc.* **2009**, *131*, 8134–8140.
- 32 Chen, X.; Zhang, L.; Zhang, L.; Wang, J.; Liu, H.; Bu, Y. *J. Phys. Chem. B* **2009**, *113*, 16681–16688.
- 33 Liu, Z.; Tan, C.; Guo, X.; Kao, Y. T.; Li, J.; Wang, L.; Sancar, A.; Zhong, D. *PNAS* **2011**, *108*, 14831–14836.
- 34 Brazard, J.; Usman, A.; Lacomat, F.; Ley, C.; Martin, M. M.; Plaza, P.; Mony, L.; Heijde, M.; Zabulon, G.; Bowler, C. *J. Am. Chem. Soc.* **2010**, *132*, 4935–4945.
- 35 Byrdin, M.; Eker, A. P. M.; Vos, M. H.; Brettel, K. *PNAS* **2003**, *100*, 8676.
- 36 Krapf, S.; Koslowski, T.; Steinbrecher, T. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9516–9525.
- 37 Woiczikowski, P. B.; Steinbrecher, T.; Kubař, T.; Elstner, M. *J. Phys. Chem. B* **2011**, *115*, 9846–9863.
- 38 Aubert, C.; Vos, M. H.; Mathis, P.; Eker, A. P. M.; Brettel, K. *Nature* **2000**, *405*, 586–590.
- 39 Rajski, S. R.; Kumar, S.; Roberts, R. J.; Barton, J. K. *J. Am. Chem. Soc.* **1999**, *121*, 5615–5616.
- 40 Boon, E. M.; Livingston, A. L.; Chmiel, N. H.; David, S. S.; Barton, J. K. *PNAS* **2003**, *100*, 12543–12547.
- 41 Yavin, E.; Boal, A. K.; Stemp, E. D. A.; Boon, E. M.; Livingston, A. L.; O'Shea, V. L.; David, S. S.; Barton, J. K. *PNAS* **2005**, *102*, 3546–3551.
- 42 Leea, P. E.; Dempleb, B.; Barton, J. K. *PNAS* **2009**, *106*, 13164.
- 43 Milligan, J. R.; Tran, N. Q.; Ly, A.; Ward, J. F. *Biochemistry* **2004**, *43*, 5102–5108.
- 44 Tong, G. S. M.; Kurnikov, I. V.; Beratan, D. N. *J. Phys. Chem. B* **2002**, *106*, 2381–2392.
- 45 Beljonne, D.; Pourtois, G.; Ratner, M. A.; Bredas, J. L. *J. Am. Chem. Soc.* **2003**, *125*, 14510–14517.
- 46 Weatherly, S. C.; Yang, I. V.; Throp, H. H. *J. Am. Chem. Soc.* **2001**, *123*, 1236–1237.

- 47 Carra, C.; Iordanova, N.; Hemmes-Schiffer, S. *J. Phys. Chem. B* **2002**, *106*, 8415–8421.
- 48 Yue, D. J. E.; Lim, K. W.; Phan, A. T. *J. Am. Chem. Soc.* **2011**, *133*, 11462–11465.
- 49 Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. *Nucl. Acids Res.* **2009**, *37*, 1713–1725.
- 50 Egli, M.; Pallan, P. S. *Curr. Opin. Struc. Biol.* **2010**, *20*, 262–275.
- 51 Takahama, K.; Sugimoto, C.; Arai, S.; Kurokawa, R.; Oyoshi, T. *Biochemistry* **2011**, *50*, 5369–5378.
- 52 Hu, L.; Lim, K. W.; Bouaziz, S.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 16824–16831.
- 53 Guédin, A.; Gros, J.; Alberti, P.; Mergny, J. L. *Nucl. Acids Res.* **2010**, *38*, 7858–7868.
- 54 Lusvarghi, S.; Murphy, C. T.; Roy, S.; Tanious, F. A.; Sacui, I.; Wilson, W. D.; Ly, D. H.; Armitage, B. A. *J. Am. Chem. Soc.* **2009**, *131*, 18415–18424.
- 55 Lim, K. W.; Lacroix, L.; Yue, D. J. E.; Lim, J. K. C.; Lim, J. M. W.; Phan, A. T. *J. Am. Chem. Soc.* **2010**, *132*, 12331–12342.
- 56 Zhang, Z.; Dai, J.; Veliath, E.; Jones, R. A.; Yang, D. *Nucl. Acids Res.* **2010**, *38*, 1009–1021.
- 57 Alcaro, S.; Artese, A.; Costa, G.; Distinto, S.; Ortuso, F.; Parrotta, L. *Biochimie* **2011**, *93*, 1267–1274.
- 58 Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345–353.
- 59 Woiczikowski, P. B.; Kubař, T.; Gutiérrez, R.; Cuniberti, G.; Elstner, M. *J. Chem. Phys.* **2010**, *133*, 035103.
- 60 Shlyahovsky, B.; Li, Y.; Lioubashevski, O.; Elbaz, J.; Willner, I. *ACS nano* **2009**, *3*, 1831–1843.
- 61 Wang, X.; Huang, J.; Zhou, Y.; Yan, S.; Weng, X.; Wu, X.; Deng, M.; Zhou, X. *Angew. Chem., Int. Ed.* **2010**, *49*, 5305–5309.
- 62 Huang, Y. C.; Cheng, A. K. H.; Yu, H. Z.; Sen, D. *Biochemistry* **2009**, *48*, 6794–6804.

- 63 Hagelstrom, R. T.; Blagoev, K. B.; Niedernhofer, L. J.; Goodwin, E. H.; Bailey, S. M. *PNAS* **2010**, *107*, 15768–15773.
- 64 Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucl. Acids Res.* **2001**, *29*, 2860.
- 65 Morozova, N.; Allers, J.; Myers, J.; Shamoo, Y. *Bioinformatics* **2006**, *22*, 2746.
- 66 Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265–322.
- 67 Barbara, P. F.; Meyer, T. J.; Ratner, M. A. *J. Phys. Chem.* **1996**, *100*, 13148–13168.
- 68 Jortner, J. *J. Chem. Phys.* **1976**, *64*, 4860.
- 69 Jortner, J.; Bixon, M. *Ber. Bunsen-Ges. Phys. Chem.* **1995**, *99*, 296.
- 70 Voityuk, A. A.; Jortner, J.; Bixon, M.; Rösch, N. *J. Chem. Phys.* **2001**, *114*, 5614–5620.
- 71 Mühlbacher, L.; Egger, R. *J. Chem. Phys.* **2003**, *118*, 179.
- 72 Voityuk, A. A.; Rösch, N.; Bixon, M.; Jortner, J. *J. Phys. Chem. B* **2000**, *104*, 9740–9745.
- 73 Barnett, R. N.; Cleveland, C. L.; Joy, A.; Landman, U.; Schuster, G. B. *Science* **2001**, *294*, 567–571.
- 74 Barnett, R. N.; Cleveland, C. L.; Landman, U.; Boone, E.; Kanvah, S.; Schuster, G. B. *J. Phys. Chem. A* **2003**, *107*, 3525–3537.
- 75 Rak, J.; Voityuk, A. A.; Marquez, A.; Rosch, N. *J. Phys. Chem. B* **2002**, *106*, 7919–7926.
- 76 Voityuk, A. A.; Rösch, N. *J. Phys. Chem. B* **2002**, *106*, 3013–3018.
- 77 Voityuk, A. A. *J. Chem. Phys.* **2005**, *123*, 034903–5.
- 78 Hoffman, B. M.; Ratner, M. A. *J. Am. Chem. Soc.* **1987**, *109*, 6237–6243.
- 79 Hoffman, B. M.; Ratner, M. A.; Wallin, S. A. *Electron Transfer in Inorganic, Organic, and Biological Systems*; ACS Publications: Washington, DC, 1991; pp 125–146.
- 80 Voityuk, A. A.; Siritwong, K.; Rösch, N. *PCCP* **2001**, *3*, 5421–5425.

- 81 Troisi, A.; Orlandi, G. *J. Phys. Chem. B* **2002**, *106*, 2093–2101.
- 82 Tavernier, H. L.; Fayer, M. D. *J. Phys. Chem. B* **2000**, *104*, 11541–11550.
- 83 Siriwong, K.; Voityuk, A. A.; Newton, M. D.; Rösch, N. *J. Phys. Chem. B* **2003**, *107*, 2595–2601.
- 84 LeBard, D. N.; Lilichenko, M.; Matyushov, D. V.; Berlin, Y. A.; Ratner, M. A. *J. Phys. Chem. B* **2003**, *107*, 14509–14520.
- 85 Matyushov, D. V.; Voth, G. A. *J. Chem. Phys.* **2000**, *113*, 5413–5424.
- 86 Kurnikov, I. V.; Tong, G. S. M.; Madrid, M.; Beratan, D. N. *J. Phys. Chem. B* **2001**, *106*, 7–10.
- 87 Basko, D. M.; Conwell, E. M. *Phys. Rev. Lett.* **2002**, *88*, 098102.
- 88 Schuster, G. B. *Acc. Chem. Res.* **2000**, *33*, 253–260.
- 89 O’Neil, M. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 11471–11483.
- 90 Voityuk, A. A. *J. Phys. Chem. B* **2005**, *109*, 17917–17921.
- 91 Conwell, E. M.; Rakhmanova, S. V. *PNAS* **2000**, *97*, 4556–4560.
- 92 Schuster, G. B.; Landman, U. *Top. Curr. Chem.* **2004**, *236*, 139–161.
- 93 Conwell, E. M.; Park, J. H.; Choi, H. Y. *J. Phys. Chem. B* **2005**, *109*, 9760–9763.
- 94 Schrödinger, E. *Ann. Phys.* **1926**, *79*, 361.
- 95 Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover, 1996.
- 96 Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular electronic-structure theory*; Wiley, 2000.
- 97 Yarkony, D. R. *Modern Electronic Structure World Scientific: Singapore*; 1995; Vol. 1.
- 98 Jensen, F. *Introduction to computational chemistry*; Wiley, 2007.
- 99 Newton, M. D. *Chem. Rev.* **1991**, *91*, 767–792.
- 100 Cave, R. J.; Newton, M. D. *Chem. Phys. Lett.* **1996**, *249*, 15–19.

- 101 Voityuk, A. A. *Chem. Phys. Lett.* **2007**, *439*, 162–165.
- 102 Cave, R. J.; Newton, M. D. *J. Chem. Phys.* **1997**, *106*, 9213–9226.
- 103 Voityuk, A. A.; Rösch, N. *J. Chem. Phys.* **2002**, *117*, 5607–5616.
- 104 Bixon, M.; Jortner, J.; Verhoeven, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 7349.
- 105 Murrell, J. N. *J. Am. Chem. Soc.* **1959**, *81*, 5037.
- 106 Rust, M.; Lappe, J.; Cave, R. J. *J. Phys. Chem. A* **2002**, *106*, 3930–3940.
- 107 Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, B864–B871.
- 108 Thomas, L. H. *Proc. Cambridge Phil. Soc.* **1927**, *23*, 542.
- 109 Fermi, E. *Z. Physik* **1928**, *48*, 73.
- 110 Kohn, W.; Sham, L. J. *Phys. Rev. A* **1965**, *140*, 1133.
- 111 Laming, G. J.; Termath, V.; Handy, N. C. *J. Chem. Phys.* **1993**, *99*, 8765.
- 112 Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- 113 Perdew, J. P.; Bruke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.
- 114 Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- 115 Harris, J. *Phys. Rev. A* **1984**, *29*, 1648.
- 116 Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009–4037.
- 117 Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359–401.
- 118 Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.
- 119 Andersson, K.; Malmqvist, P. A.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
- 120 Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
- 121 Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. *Chem. Phys. Lett.* **1998**, *288*, 299–306.

- 122 Dickerson, R. E. *Nucl. Acids Res.* **1989**, *17*, 1797–1803.
- 123 Lu, X. J.; Olson, W. K. *J. Mol. Biol.* **1999**, *285*, 1563–1575.
- 124 Brettel, K.; Byrdin, M. *Curr. Opin. Struc. Biol.* **2010**, *20*, 693–701.
- 125 Mor, M.; Silva, C.; Vacondio, F.; Plazzi, P. V.; Bertoni, S.; Spadoni, G.; Diamantini, G.; Bedini, A.; Tarzia, G.; Zusso, M. *J. Pineal Res.* **2004**, *36*, 95–102.
- 126 Reiter, R. J.; AcuXa-Castroviejo, D.; Tan, D. X.; Burkhardt, S. *Ann. NY Acad. Sci.* **2006**, *939*, 200.
- 127 Reiter, R. J.; Tan, D. X.; AcXua-Castroviejo, D.; Burkhardt, S.; Karbownik, M. *Curr. Top. Biophy.* **2000**, *24*, 171.
- 128 Gürkök, G.; Coban, T.; Suzen, S. *J. Enzyme Inhib. Med. Chem.* **2009**, *24*, 506–515.
- 129 Genereux, J. C.; Boal, A. K.; Barton, J. K. *J. Am. Chem. Soc.* **2010**, *132*, 891–905.
- 130 Butchosa, C.; Simon, S.; Voityuk, A. A. *Org. Biomol. Chem.* **2010**, *8*, 1870–1875.
- 131 Butchosa, C.; Simon, S.; Voityuk, A. A. *Int. J. Quant. Chem.* **2012**, *112*, 1838–1843.
- 132 Butchosa, C.; Simon, S.; Voityuk, A. A. *Comp. Theor. Chem.* **2011**, *975*, 38–41.
- 133 Andersson, K.; et al., *MOLCAS, version 7.2*; University of Lund: Lund, Sweden, 2008.
- 134 Félix, M.; Voityuk, A. A. *Int. J. Quant. Chem.* **2011**, *111*, 191–201.
- 135 Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
- 136 Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- 137 Chai, J. D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- 138 Frisch, M. J.; et al., In *Gaussian 03*; Gaussian, I., Ed.; 2004; Vol. Rev. E.03.

- 139 Seminario, J. M. *Recent developments and applications of modern density functional theory*; Elsevier Science, 1996; Vol. 4.
- 140 Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz Jr, J. L. *Phys. Rev. Lett.* **1982**, *49*, 1691–1694.
- 141 Zhang, Y.; Yang, W. *The Journal of chemical physics* **1998**, *109*, 2604.
- 142 Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 8425.
- 143 Rutledge, L. R.; Durst, H. F.; Wetmore, S. D. *J. Chem. Theory Comput.* **2009**, *5*, 1400–1410.
- 144 Kobayashi, K.; Tagawa, S. *J. Am. Chem. Soc.* **2003**, *125*, 10213–10218.