New Computational Methods and Developments for Molecular Response Properties in Condensed Matter

Vincenzo Verdolino

General and Inorganic, Analytical and Physical Chemistry Department

University of Parma, SS.MM.FF. Faculty

A thesis submitted for the degree of

PhilosophiæDoctor (PhD) in Chemistry

2009 March

1. Supervisor: Prof. Roberto Cammi

2. Reviewer: Dott. Paolo Pelagatti

Date of the defense: March 2009

Signature from head of PhD committee:

Abstract

In this thesis we are aimed by the purpose of describe three important cases of chemical interest and show how theoretical and computational methods can be useful, even needful to explore and solve very significative applications. The commons theme of this work are represented by the well known Density Functional Theory and the Polarizable Continuum Model used for all the cases reported . As we will able to recall the first is one of the most diffused theoretical approach in computational chemistry studies whenever the electronic structure has to be solved. The second one represents a wide diffuse model which is able to describe with good accuracy molecular interactions in condensed matter. Both of them provide the most common computational configuration in organic, and oreganometallic studies.

In this thesis we are going to describe few examples which spread from the study of condensed matters under extremely high pressure to a biological application on DNA nucleobase systems, moving through some typical organometallic problems. Focusing on our issues we are going to present a new computational QM method for the study of structural properties (i.e. equilibrium geometry) of molecular systems under very high pressure. The procedure is based on the Polarizable Continuum Model , usually used to study molecular solutes under standard pressure conditions. The presented development considers two critical items: the definition of the pressure and the elaboration of an analytical code for the calculation of molecular gradient. The method has been developed at HF and DFT levels, with computational costs comparable with those of similar calculations *in vacuo*. The numerical examples regarding the equilibrium geometries and conformational energies (DFT level) of 1,3-butadiene under high pressure give an indication of the potentialities of the approach and of the problems to which it may be applied.

The next argument regards several application in the field of inorganic chemistry. We will shortly describe a computational procedure where we investigate some spectroscopical properties of some luminescence silver complexes. In this, we will describe a time dependent density functional study of several excited states. Moreover we will study the correlation between stretching vibration and Mulliken charges in some rhodium complexes. But most of this section will be dedicate to a very practical application in chemistry that is the chemical reaction path determination. Aimed by unclear experimental evidences provided by the catalytic investigation of an half sandwich Ru(II) complex, we performed a deep exploration of the mechanism through which the pre catalysts operate. Several data obtained by high-resolution MS (ESI) have been explained by a combined "Density Functional/Polarizable Continuum Model" study. The results reveal that the complexes containing PNH_2 operate through a bifunctional mechanism analogous to that proposed for diamines and amino alcohol ligands, but with some new aspects that are important to point out.

Finally an efficient computational method has been identified which uses Density Functional Theory, Polarizable Continuum Models with a modified UFF cavity and Boltzmann weighting of tautomers to predict the site-specific and global pK_a of DNA nucleobases and their oxidation products. The method has been used to evaluate the acidity of Gh and Sp, two highly mutagenic guanine oxidation products. The trend observed for the pK_a values of Gh (9.64 and 8.15) is consistent with the experimentally observed values for guanidine cation (13.7) and hydantoin (9.16). The pK_{a1}(calc) value for deprotonation of Sp cation ($Sp^+ \rightarrow$Sp) is very close to the experimentally observed pK_{a1} for 8-oxoG and is consistent with the similarity in their structures. The data suggest that the imide (N7) proton in Sp is considerably more acidic than that in Gh, possibly due to the presence of the through-space electronic effects of the carbonyl group located at C6. This difference in the acidity of Gh and Sp may be an indication of their potential toxicity and mutagenicity in vivo and remains a fertile area for experimental study.

To My Parents

Antonia and Giuseppe

Acknowledgements

Voglio ringraziare i miei genitori Mamma e Pap. A loro dedicata questa tesi di dottorato. Solo loro hanno consentito che i lunghi anni di studio potessero condurmi al completamento di tre anni trascorsi nel campo della ricerca. Sono stati tre anni estremamente importanti ed grazie a voi se ho potuto viverli. Il vostro amore mi sempre stato vitale.

Ringrazio poi naturalmente il mio Amore. Chiara mi sei sempre stata vicino, abbiamo fatto sempre tutto insieme. Sei stata la mia forza e mi sei stata di consiglio in tutti questi anni. Continuerai ad esserlo a lungo. Ti amo

Voglio senza dubbio ringraziare inoltre il mio mentore, il Professore Roberto Cammi che oltre ad avermi formato durante la didattica universitaria mi ha iniziato al mondo della Chimica Teorica durante la mia tesi di laurea e mi ha fornito tutti gli strumenti, nonch fondamentali direttive per lo svolgimento e completamento di questa tesi di dottorato. La ringrazio infinitamente e spero di poter presto tornare a lavorare con lei.

I also want to thanks Prof. H. B. Schlegel, Jason Sonnenberg, his beautiful family and Barbara Munk who gave me a wonderful hospitality in the USA. With them I had beautiful work and spare time. Thank you.

Contents

List of Figures

List of Tables

List of Equations

LIST OF EQUATIONS

1

Introduction

1.1 Chemistry and Computation

For a long time, experiments in the chemistry lab have been carried out following pure empirical procedure and without detailed calculations and predictions. However, in the last few ten years most organic, inorganic and physical chemists started to use results of Computational Chemistry to explain several unclear researching fields and to predict *phenomena* and behavior of matter (1). A wide range of theory and methodology have been developed by the years so far and some of them were oriented to prescind by experimental data and references. Such kind of approaches are called *ab initio* methods and they represent the most tempting and ambitious field of chemical and physical theory.

Density functional theory (DFT) has been pretty recently introduced among the theoretical methods and this meant a critical change and improvements in the computational chemistry applications (2). Density functional calculations is nowadays a strong and powerful tool to help to decide between alternative synthetic routes and to provide guidance on structure assignments. Scientists employed in theoretical developments are focusing on different areas and many of them work to make progress which can be made in the near future on problems that will benefit their experimental colleagues in chemistry and other disciplines. One of the most important help provided by computation to the experimental work is the discrimination among several reaction pathways. Frequently this represents a powerful strategy to force complicate reactions into the desired way or anyway a tool for investigation and definitely save job, time and money. The first important step to define a complete reaction path is the construction of the potential energy surface, which describes the energy of the molecule or them assembly as

a function of the atomic positions. In this context it is clear that moving towards large systems the computing time could increase very quickly and a principal rule is played by development of efficient methodology and computing technology.

It is well known that predicting reactions between isolated small molecules in the gas phase context theoretical chemist can reach the highest level of quantum mechanical detail. Reactions involving three atoms have been modeled successfully in this way, but even four- or six-atom reactions remain challenging. Although DFT has been the workhorse of most computational studies, it is not reliable in certain classes of problems; for example, it often underestimates transition-state barriers in reactions or band gaps of materials. Many other problems in DFT approximations are related to de localization and static correlation errors through a framework that makes use of fractional charges and fractional spins. In the following chapters we will able to treat some of the advantage and drawback of the DFT methods.

Practical chemistry often involves molecular interactions in multiple phases beyond gaseous collisions. Several progress in modeling molecule/surface interactions, which play a central role in heterogeneous catalysis, have been established recently and they represent one of the most interesting goal of theory.

Theoretical chemistry is now commonly used to address complex problems in biochemistry and materials science. Recent successes has been achieved in simulating protein folding, a problem long hindered by the computational intractability of the immense number of accessible configurations. Recent progress in ab initio and DFT methods has facilitated purely theoretical explorations of features ranging from mechanical properties to corrosion behavior. Electronic excitation remains a challenging frontier. The principal theme which is found every time we went to explore physics with theoretical tools is the complementary role of theory itself and experiment in exploring chemical questions. Each approach nourishes the other, presenting fresh challenges.

2

Theoretical Background

2.1 Quantum Mechanics and Computational Chemistry

Over the past three decades, ab initio quantum chemistry has become an essential and powerful tool in the study of atoms and molecules and, increasingly, in modeling complex systems such as those arising in biology and materials science. Is well known that a complete description of a system is obtainable with the solution of the Scrodinger's eigenvalue equation. In particular, in chemistry, the Scrodinger problem reduces to solve the electronic structure of atoms and more in general molecules. The ability to obtain goodenough solutions to the electronic Schrodinger equation for systems containing tens, or even hundreds, of atoms has given the possibility to theoretical chemist to investigate and solve in many cases, important and tricky problems in a wide range of fields. In its exact form, the electronic Schrodinger equation is a many-body problem, whose computational complexity grows exponentially with the number of electrons, and hence, a solution is intractable. HartreeFock (HF) theory is a mean field approach to solve the many-electron problem using a Self Consistent (SCF) procedure. The main lack of the HF theory consists in the absence of explicit electronic correlation term which represents an important contribution to the total energy. For this reason the HF produces gives reasonable results for many atomic and molecular properties but is incapable of providing a robust description of reactive chemical events in which electron correlation has a major role. Thus, a key problem has been the development of treatments of electron correlation that exhibit a tractable scaling in computational effort with the size of the system.

Another important improvement into chemical description is played by theory that treat solid and solution environment. This is one of the most important goal to reach with the purpose

of describe in a realistic way a complex chemical system. The treatment of large, condensed phase systems (e.g., proteins in aqueous solution) entirely by ab initio methods is extremely expensive computationally. However, it is often the case that a relatively small region of the system can be modeled at the ab initio quantum chemical level, whereas the remainder can be treated more approximately (e.g., by means of molecular mechanics (MM) or continuum solvation models).

The ultimate goal of theory and computational chemistry is that of achieving strong partnership with experiment as both an explanatory and predictive methodology. In this way the theoretical approach will be able to provide a very useful tool to chemists in the exclusion of such no productive directions and explain unclear behavior which could open new interesting research. Continued improvements in the theory and computing performance (including coding and computer technology), ensure that very important progress will continue to be made in the next years.

Two significative approaches to solution of the electronic Schrodinger equation have been made over the past 50 years: the Wavefunction-based and Density Functional approaches. Wavefunction-based approaches expand the electronic wavefunction as a sum of Slater determinants, the orbitals and coefficients of which are optimized by various numerical procedures. HF theory is a particular case of this type, involving optimization of a single determinant which represents the grounds electronic state. For this reason however, as mentioned above, its usefulness is limited because of complete neglect of electron correlation. More accurate wavefunction methods has been developed, such as Coupled cluster (CC),Mller-Plesset (MPn) and Multireference perturbation (CASPTn) methods. Density functional theory methods are based on the Hohenberg Kohn theorem (3) and involve a functional description of the electron density to obtain the expression of the total free energy of the system. Because the electron density depends on only three coordinates (as opposed to the 3N coordinates of N electrons), the computational effort required to solve the equations of DFT is comparable with that required for HF theory, thus rendering DFT highly attractive from the point of view of computational implementation. However, the correct functional of the energy is unknown and has to be constructed by approximation.

Earlier functionals, based principally on behavior of the electron gas model (4), were lacking in the accuracy required for chemical applications. Breakthroughs over the past two decades (5; 6; 7; 8; 9) have led to the development of functionals capable of remarkable accuracy and applicability across the periodic table and complex molecules. It is important to

note that there remain limitations in DFT methods that implies the usage of different theory for many application field. At present, there are two principal classes of functionals that have been extensively tested in large scale applications as well as small molecule benchmarks: gradient-corrected (6; 8; 9) (e.g., BLYP; ref. (6)), and hybrid (6; 8) (e.g., B3LYP; ref. (6), functionals. Gradient-corrected functionals begin with the local-density approximation but add terms involving the gradient of the electron density. Hybrid functionals also incorporate gradient corrections but add an empirically fitted admixture of exact Hartree Fock exchange.

CC methods are the most computationally expensive but also, in principle, the most accurate of the approaches that we are briefly listing. An exponential ansatz is used in CC for the coupling of correlated electron pairs and highly excited determinants are included into the wavefunction setting up. In this theory the only one limit is the computational cost with the increasing of the number of single, double and triple, etc., excitation used to build the wave function. Has been demonstrated that the CCSD(T) approach (in which triple excitations are treated perturbatively) provides the best trade-off of accuracy and efficiency (10). The formal scaling with number of electrons N of CCSD(T) is N^7. Thus, as anticipated, calculations are limited to small- to medium-sized molecules, and parallel supercomputers are required. On the other hand the level of accuracy reached from a benchmark CCSD(T) calculation for atomization energies (and other less demanding thermo chemical quantities) of small organic molecules is on the order of a few tenths of 1 kcal/mol (1 cal = 4.18 J) (11). Equilibrium geometry parameters are typically highly accurate (12). As the level of accuracy achieved in geometry optimization is not so determining factor for energetic accuracy it is common to use a less expensive methods to compute structures. Than, with a reasonable geometry close to the equilibrium minimum, it is possible to perform an high accuracy calculation with CC methods.

Turning next to MP2 methods, the computational cost is reduced considerably by two or more orders. Numerical methods developed in the past ten years has been able to decrease the formal scaling N^5, both formally and asymptotically enabling the investigation of systems containing a few hundred atoms. It is important to understand when a MP2 methods is better than a CC approach: the results of MP2 methods for properties involving making or breaking bond or in general electron pairs are worst than to that of CCSD(T) and also not as good as the best DFT methods, although recent studies have come closer to DFT benchmarks (13). However, for many cases where the study is focused on conformational and no bonded interaction, MP2 methods do extremely well and the accuracy become ≈ 0.3 kcal/mol if one extrapolates to the basis set limit (14; 15; 16).

Multireference perturbation methods, such as CASPT2 (17), are more difficult to use than either CCSD(T) or MP2 in that they require a careful definition of the active space of electrons to be treated at the multireference level to be effective. In this approach all the electrons not included in the active space are treated perturbatively. The choice of the active space dimensions, represents the critical step of the multireference methods because of the ratio accuracy/computational-time. The computational cost, in fact, increases exponentially with the size of the active space hence very quickly. In some cases we can find few problems to treat the system with CCSD(T) methods. In particular severe difficulties could be found because of the fact that the electronic wavefunction is no longer dominated by a single (HartreeFock) determinant. The most important case is when a complete potential energy curve during the process of bond breaking is under study (18).

Coming back to DFT, both gradient correct and hybrid methods achieve excellent results for equilibrium geometries (19; 20) across the periodic table. In fact, the DFT methods in general, represent an excellent way to get geometry very close to the equilibrium and useful for further studies. This is no longer true for properties which are strongly dependent on the electronic correlation. For atomization energies, there is a clear difference in performance between different DFT methods; for example, the BLYP functional has an average error of 7.09 kcal/mol for the G2 set of 148 small molecules, whereas the B3LYP functional has an average error of 3.11 kcal/mol for the same data set (21). The three subjects under study in this thesis are related by some features that represent the linkage between all of them. The first important one is the quantum chemical method used to solve the electronic structure: the DFT approach. In the Sections 3.1 and 4 we are going to show some developments and results related to the effects that pressure has on several molecular properties. It has been a critical point the choice of DFT approach because of the behavior of the functional B3LYP within high electron repulsion interaction.

Particular attention must be paid in transition-metal-containing systems. In Sections 3.2 and 4.2 we will study three chemical systems containing metals from different point of view. In these cases the experimental data are less reliable, and there are many different types of bonding to consider. However the DFT-B3LYP shows satisfactory results with average errors for bond energies of small transition metal complexes in the range of 35 kcal/mol (22). Thus, hybrid DFT provides a quantitative means of investigating reactive chemistry in large systems, albeit with large errors for a small number of outliers. Moreover in Section 4.2.1 we use the time-dependent version of DFT (TDDFT) which, in many cases, provides reasonable results

for excited states (23) and is considerably less expensive than alternatives such as CASPT2- and CCSD(T)-based excited-state methods. All these well known features of DFT justify our choice in Section 3.2.2 where we study a complex chemical system involved into a catalytic hydrogen transfer. In the related Resultes Chapter-Section 4.2.2 we will show a typical case where the DFT-B3LYP approach is a good way to describe a complex chemical mechanism and, moreover, a strong way to obtain related details hardly to get with others methods. In the same Mechanistic Results Section 4.2.2, we meet a critical point when a theoretical chemist uses DFT methods to describe chemical and physical interaction. As it will be showed, our catalyst is involved in an hydrogen transfer between two organic molecules. Moreover two possible ways could be follow to reach the desired products which differs by chirality. Our studies bring to the conclusion that a no bonded and dispersive interaction is strongly involved in this critical step. However, the ability of DFT to describe dispersion (van der Waals) interaction has not yet been established (14). In our case, like many others similar, MP2 represents a qualitatively superior alternative, which can itself be applied to large systems at a quantitatively (but not qualitatively) larger computational cost.

One more DFT study has been developed and reported in this thesis, and it is described in Section 3.3 and its related resultes in Section 4.3. With this last application we will show the efficiency and accuracy of DFT in prediction of not only molecular properties but also very important and useful thermodynamic reaction quantities .

One last important theoretical improvement has been used in this thesis and it represents the second important connection between the three mean arguments under study. This is the theory and the computational tool used to describe the physical environment surrounding our objects and reactions. A complete description at quantum level of a molecular system interacting with its solvation environment, as well as any other very large ensemble of particles, represent, at this time, an inaccessible investigation field.

Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) methods (24) provide a means of combining quantum chemical and molecular mechanics models in a way that preserves the integrity of the overall energy function. The key issue is the interface between the QM and MM regions. In principle two different situation can be occur: the main interaction between the QM and the MM region has an electrostatic nature or the two portion are connected by one or more covalent bond. If the QM and MM regions are not covalently linked, then the problem is relatively straightforward; introduction of van der Waals parameters onto the QM atoms, and optimization of these parameters to fit fully QM hydrogen bonding energies , will

yield a model that does a reasonable job of reproducing the fully QM results over the entire phase space. However, when there are covalent connections between regions , definition of the interface requires more sophisticated techniques. The QM/MM approach is ideally suited to studying reactions in proteins, disordered solids, large molecules, and other systems that have well defined structures and are heterogeneous in character.

Description of a solution phase requires a different approach, in fact for processes in solution, it is necessary to average extensively over the positions of the solvent molecules to properly take into account environmental effects. This is translated into a limitation for the QM/MM approaches which are possible in principle but computationally very expensive. An alternative that delivers reasonable accuracy at a much lower computational cost is the use of solvation continuum models . The coupling of continuum models with quantum chemical calculations using the self-consistent reaction field (SCRF) approaches (25; 26) has been implemented over the past decade in a number of widely available ab initio quantum chemistry programs (e.g., GAUSSIAN, JAGUAR, and QCHEM). As we will focus in the Section 2.3, alternating cycles of continuum Poisson Boltzmann (27) (or approximations of such cycles; Ref. (28)) and quantum chemical computations, the interacting reaction field of the solvent and charge distribution of the solute are iterated to convergence. Not only single point evaluation are available, even better full geometry optimization of structures in solution is also possible by using analytical gradient methods (29; 30). In the Section 3.1 we show an example of application of analytical gradient theory developing the first derivative expression of the Repulsion free energy included in the SCRF method used to describe molecular system under high pressure. It is important to point out that to achieve accurate results for solvation- free energies, parameters of the continuum model must be fit to small molecule experimental data (28; 31). SCRF-based methods can also be used effectively for pK_a prediction (32) with several problems that will be discussed and solved in part in the Sections 3.3 and 4.3.

We can summarize these firsts notions saying that for small molecules in the gas phase, CCSD(T) and related methods are the best approach , perhaps combined with less expensive methods to carry out geometry optimization and/or obtain basis set convergence. In such cases the CCSD(T) calculations could be more accurate than experiments. For investigation of reactive chemistry in medium to large systems, DFT is at present the best approach. MP2 methods are generally useful for studying non bonded interactions in both small and large systems and in molecular mechanics force fields development. DFT can be used also in the investigation of conformational and hydrogen bond interactions, provided that the accuracy limitations are

understood. Multireference perturbation methods are typically applied when there are particular difficulties associated with "near degeneracy", or for properties such as excited states for which specific multireference approaches (such as CASPT2) have proven to be superior to DFT in a wider subset of cases. The overwhelming majority of large-scale material science and biological applications have been performed with DFT; this state of affairs is mandated by the large size of the systems that are being considered, the need in some cases for periodic boundary conditions, and the availability of QM/MM and SCRF methods principally for DFT based approaches.

2.2 Density Functional Theory

As we introduced in the previous section, DFT is a QM theory used in physics and chemistry to investigate the electronic structure of many-body systems. As is well known DFT was given a formal footing by the two theorems introduced by Hohenberg and Kohn (HK) in 1965 (3). In this theory, all the molecular properties are described solving a many-electron problem by using functionals (i.e. functions of another function, which in this case is the spatially dependent electron density). In particular Hohenberg and Kohn proved that each contribution to the total energy can be expressed as a functional of the total electron density

$$\rho(r) = \sum_{\gamma=\alpha}^{\beta} \rho^{\gamma}(r) = \sum_{i\in\alpha} \psi_i^2(r) + \sum_{i\in\beta} \psi_i^2(r) \tag{2.1}$$

(α and β are the two possible spin states), by the following energy functional relation:

$$E_V[\rho] = F_{HK} + \int \rho(r)V(r)dr \tag{2.2}$$

In this section we briefly show the main aspect of the DFT theory recalling the two fundamental theorems and the principal development of theory.

The first HK theorem demonstrates that the ground state properties of a many-electron system are uniquely determined by an electron density that depends on only 3 spatial coordinates. Hence through the use of functionals of the electron density, the 3N variables problem of a N body-system reduces to a 3 spatial coordinates issue. The second HK theorem defines

an energy functional for the system and proves that the correct ground state electron density minimizes this energy functional.

In the DFT theory the interaction between bodies deals with an effective potential which include the external and the *electron-electron* interaction terms. A set of one-particle eigenvalue equations developed by Kohn and Sham (KS) (33; 34) provide the explicit solution to the DFT problem:

$$H^{\gamma}_{KS}\phi_{i\sigma}(r_1) = \left(-\frac{1}{2}\Delta + \sum_A \frac{Z_A}{|R_A - r_1|} + \int \frac{\rho^{\gamma}(r_2)}{|r_1 - r_2|} + V_{XC}(r)\phi_{i\sigma}(r_1)\right) = \epsilon_{i\sigma}\phi_{i\sigma}(r_1) \qquad (2.3)$$

The first term in the KS Hamiltonian represent the kinetic energy of the N non-interacting electrons, and the second and the third represent the classical nucleus-electron and electron-electron Coulomb interactions. The last term V_{XC} is the exchange-correlation potential which include all no classical exchange and correlation interactions. In this the no classical effects on the electron kinetic energy are also take into account. The V_{XC} is defined as the partial derivative of the exchange-correlation energy (E_{XC}) with respect to the electron density ρ. Considering the equations 2.1 and 2.3 it is obvious that the KS equations are a pseudo-eigenvalue problem and has to be solved iteratively as in the HF scheme, hence through a self-consistent field procedure. On the other hand the HF and KS solutions are substantially different and, in principle, non-physically comparable: it is possible to demonstrate that in general the energy of KS orbitals doesn't have physical meaning, in fact the HK theorem refers only to the exact ground state one-electron densities, and hence, it cannot be regarded as the basis upon which to exact the KS ansatz for the exchange-only self-consistent field. Moreover it is not a good method for the calculation of the excitation energy. For this argumentation in the KS formulation, the eigenvalues are just considered as auxiliary quantities for computing of corrected total energy and one-electron density. The original eigenvalues are not related to any physical meaning except for the highest occupied orbital (HOMO) which determinate the first ionization energy changed in sign. In other words the density functional approach leads to an exact solution of the N-electron problem once the true energy functional exchange-correlation (E_{XC}) is available. This represents the tricky point of DFT application: no one knows the correct functional E_{XC} for atoms and molecules. Therefore approximation have to be made to provide expressions for E_{XC}, as reported in the next paragraphs.

2.2.1 Local Density Approximation

The research and developments in the DFT framework are essentially related to modeling the E_{XC} potential. As we mentioned before the simplest approximation is the local-density approximation (LDA), based upon exact exchange energy for a uniform electron gas. The Thomas-Fermi model provides the principal tool useful to obtain the exchange energy within fits to the correlation energy for a uniform electron gas. In this approximation the energy functional is a pure density functional and reads,

$$E_{XC}^{LDA} = \int \rho(r)\epsilon_{XC}[\rho(r)]dr \tag{2.4}$$

where ϵ_{XC} is the equivalent of E_{XC} per single particle. This last term can be divided in two associated to exchange and correlation. It is possible to demonstrate (35) that the exchange contribution is

$$E_{X}^{LDA} = -C_X \int \rho(r)^{\frac{4}{3}} dr \tag{2.5}$$

whereas the correlation term is provided by Monte Carlo simulations (36). The LDA has good performance describing molecular bond but fall in energy estimation: usually total exchange energies are underestimated by up to 10% while correlation energies are overestimated by up to a factor of 2. The erroneous evaluation of these two contributions is due to the approximation that describe the electron cloud like to an uniform electron gas, but it is always verified that this is far from the reality.

2.2.2 Nonlocal Corrections

An important contribution to the application of DFT to the chemical systems, comes from implementation of nonlocal correlation functionals. A number of increasingly refined mathematical expression for these functionals were introduced by Becke since 1983 (37; 38; 39). The most successful introduce additional functionals which include the gradient of the density charge (i.e. $\nabla\rho(r)$). This is a correction to the LDA:

$$E_{XC}[\rho] = \int \rho\epsilon_{XC}(\rho)dr \cdot F(\rho, \nabla\rho) \tag{2.6}$$

where the $F(\rho, \nabla\rho)$ functional is the nonlocal correction. The most commonly used nonlocal functionals are combination of these and they are BP, BLYP and BPW.

2.2.3 Hybrid Functionals

Cause of the not satisfactory treatment of the correlation energy Becke (40; 41) developed a different approach, introducing into the "wrong" DFT functional the "exact" HF exchange. This new methods gave start to several so called hybrid functionals which allowed to achieve an improved accuracy, reducing the error in the atomization energy for a set of molecules to less than 3 Kcal/mol (gradient correction only performs accuracy levels up to 6 Kcal/mol). The hybrid functionals, as anticipated, are combination of gradient-corrected DFT functionals and HF exchange contribution where the weight factors for each of them comes from a semi empirical fitting of well established experimental values. The most commonly used function is the 3-parameter functional due to Becke and include also nonlocal corrections:

$$E_{XC} = 0.2E_X^{HF} + 0.8E_X^{LDA} + 0.72E_X^B + 1.0E_C^{LDA} + 0.81E_C^{NL} \tag{2.7}$$

where the upper label B stand for the standard Becke gradient correction to the exchange and the NL ones represent a generic non-local correction to the correlation. The most popular combination is the B3LYP functional and it will be used in all the applications here described.

2.2.4 Transition metal and DFT

For the first and third issues of this thesis (see sections 3.1 and 3.3) the choice of a DFT approach is due just to make the investigation easy for very complex and new field of research. Otherwise the mechanistic exploration of a complete chemical reaction assisted by a transition metal (TM) complexes (i.e. Section 3.2.2), or a typical spectroscopical properties of inorganic compound, need that more complete argumentations must be reported. Because of their more complex electronic structure (often open shell configuration), it is important to understand if DFT methods are suitable for TM systems. In particular in such cases the most critical point is to describe exchange-correlation energies and the fact that many of these are not well described by a single Slater determinant. As we know in HF-based methods, exchange is described exactly while the correlation problem can be solved only approximately. On the other hand,

DFT methods solve approximately both of them albeit implicitly included in the exchange-correlation functional. In the absolute value the energies arise from DFT calculation show very large errors but, many times, in chemical reaction study, chemists are more interested into the relative energies and these errors cancel to large extent. This, in particular, explain the remarkable accuracy of even local DFT functionals for complex molecules. More attention must be paid if during a coordinate reaction some bonds break or a functional group changes drastically its electronic structure. This is a very frequent situation in TM system reactions. In such cases exchange and correlation variations occurs differently and DFT methods fall specially for open shell TM. However, wavefunction-based methods are facing similar difficulties in such situations and DFT is competitive also in these areas. As mentioned before, the most important advantage is represented by low computational cost (N^3 scaling factor compared with N^5 in HF-based methods for N basis functions). In other word, given a constant number of basis functions, gradient-corrected DFT or hybrid functional implementations require a computational demand which lies between HF and MP2 calculations.

DFT theory is, jointly with the Solvation Continuum Models, the common linkage between the three arguments of this thesis and we are going to conclude this brief reminding some important, but often ignored, limitation and advantage of this method. First the HK theorem applies only to ground states and a proper treatment of exited states is still an important issue in DFT. Moreover, the unrestricted variation of the KS approach for open shell systems suffers of *spin contamination*, although it is lower than in the HF methods. It is important to keep in mind that this is an unsolved question what relevance the expectation value $\langle S^2 \rangle$ has for the unknown wavefunction of the real system. A disadvantage of more practical relevance is that there is no well-defined hierarchy of functionals in DFT theory, hence all the calculations have a certain degree of trial and error and it is often not clear why one particular functional works better than another. Finally two much more technical observation must be reported. One is related to the SCF convergence which is often slow and in some cases it requires fine tuning. The other one is an important problem which frequently occurs in TM systems and it is due to a larger spatial charge variation much more the accuracy which depends on the grid mash used in the numerical integration. The price to be paid for a batter and more fine grid is of course a heavier computational demand.

2.3 Solvation Continuum Models

One of the most important tool provided by theory which enabled efficient and accurate comparison with the experimental data has be done by the modeling of liquids and solution. This is a tricky field of research and among the different approach proposed we can distinguish two principal schemes: models where the ensemble of solute and solvent are treated at the same level of theory and models where the attention is focused on the solute molecule or a very small number of molecules and the solvent description is treated at a lower level and only average described. In this section and in the remain of this thesis we will interested in these second type of approaches which are all connected by a common aspect, namely the use of a mean-field description of the chemical environment that surround and solvate the focused solute molecule . In this view the solvent is described by a continuum where take place a cavity of proper shape and dimension in which the solute is introduced. The continuum represents a dielectric making all the interaction that a solvent distribution of molecules have with the solute.

The wide diffusion and development of the solvation continuum models brought to techniques and methodology suitable for the study of small to very large molecules both for geometry optimization and energy calculation including reactivity path. This has been possible by the increase in the realism of the model and the introduction and coupling with quantum-mechanical methods. This last point is the basis that allowed the explanation of several chemical and physical phenomena not tractable with classical methods. We briefly report some technical issue related to the solvation continuum model formalism, starting from the quantistic operator $\hat{H}^{tot}(r_M)$ which represents the total energy of the interaction between solute and solvent molecules.

$$\hat{H}^{tot}(r_M) = \hat{H}^{M}(r_M) + \hat{H}^{MS}(r_M) \tag{2.8}$$

where $\hat{H}^{M}(r_M)$ is the solute molecule M Hamiltonian and $\hat{H}^{MS}(r_M)$ is the interaction between M and the continuum represent the solvent. The principal observation that we have to point out is the absence of the solvent coordinate (r_S) that represent the most important difference between continuum and discrete models. This last considers explicitly each interaction between the solute (frequently just one molecule cause of the extremely dilute solution) with a finite number of solvent molecules. We will focus on the solvation continuum models. In spite of the two different members in the Eq. 2.8, common computational codes, treats the $\hat{H}^{tot}$ as a

whole because it is an effective Hamiltonian and its structure is very similar to that of *in vacuo* calculation in the HF or DFT approach.

The $\hat{H}^{MS}(r_M)$ include several terms, one for each physical interaction between solute and solvent molecules. These contributions are provided by different operators that could be symbolically indicated with $\hat{Q}_x(\overrightarrow{r}, \overrightarrow{r'})$ where $\overrightarrow{r}$ is the position vector and x stands for one interaction. In the next subsection we are going to describe the most common used model in the framework of continuum solvation and it also will be the only one model used to treat the chemical systems under study in this thesis.

2.3.1 The Polarizable Continuum Model

In the Polarizable Continuum Model (PCM) the interaction between solute and solvent is described by four operator $\hat{Q}_x$ each having a physical nature. The solvation process is accompanied by a free energy exchange with the environment, and this value is obtained by summation of the free energy contribution of each interaction term. Including the thermal motions of the molecular framework in solution, the formal equation that provide the total free energy in solution is:

$$G(M) = G_{cav} + G_{el} + G_{dis} + G_{rep} + G_{tm} \tag{2.9}$$

Computationally speaking the formation of each term has the name "Charging Process" (CP). Each CP is based on several iteration finalized to minimize the residual coupling between an appropriate parameter and a potential function. In the PCM formalism only the first term is evaluated separately whereas the others three CP are unified and described by Schrodinger solution. The CP related to G_{cav} makes an empty cavity inside the continuum with a proper shape and size to host the solute. Only the nuclear relative position and atomic parameters are used in this route and no electronic properties of solute are involved. The calculations are performed at a given temperature T and solvent density ρ. This term represents the reversible work spent to build the cavity against the cohesion liquid forces and it could be computed with several methods (42; 43). In the actual PCM codes the standard methods is that proposed by Pierotti (44) in according to a suggestion given by Claverie (45). The parameter characterizing the solvent cavity is the solvent equivalent radius and the expression for the free energy term is analytical for spherical shape cavity and semi-analytical for the cavities built with atomic solute

spheres. In the actual PCM version the $\hat{Q}_{cav}$ operator is not included into the Hamiltonian 2.8 because in the BO approximation, for a fixed geometry this term doesn't change.

The Section 3.1 is dedicated to the effects that pressure induce on the molecular properties. We will show that acting on the equivalent radius we will able to induce significant change on the cavity size due to increasing of pressure.

What about concern the G_{el} contribution there are two different type of approach which lead to the same results but one of these is more expensive than the other. The electrostatic interaction depends on the total electron density of the focused molecule and the electric polarization of the medium. The contribution could be computed starting from non-interacting nuclei and electrons that will compose the molecule with just one CP, or computing every time the charging starting from an *in vacuo* molecular geometry and electron density and than switching on the interaction solute-solvent. This last is computationally expensive and it is required only in that cases the explicit value of solvation free energy is needed.

In the PCM formalism the medium response function for this term is the polarization $\overrightarrow{P}$:

$$\overrightarrow{P} = \frac{\epsilon - 1}{4\pi}\overrightarrow{E} \tag{2.10}$$

where $\overrightarrow{E}$ is the electric field generated by the apparent charges spread on the cavity surface, and ϵ is the dielectric constant.

To solve the electrostatic problem for the PCM it is required to solve the simple Poisson's equation $\nabla^2\varphi = -\frac{\rho}{\epsilon}$. If the medium has homogeneous and isotropic dielectric constant and the solute molecule (M), inside the cavity (C), is surrounded by an infinite and continuum dielectric medium characterized by the permeattivity ϵ, than an electrostatic field is generated and it is described by the following Poisson's equations:

$$\begin{aligned} -\Delta V &= 4\pi\rho_M && inside\ C \\ -\epsilon\Delta V &= 0 && outside\ C \\ ([V]) &= 0 && on\ the\ surface\ of\ C \\ ([\partial V]) &= 0 && on\ the\ surface\ of\ C \end{aligned} \tag{2.11}$$

where V is the electrostatic potential generated by the charge distribution ρ_M inside the cavity. The two boundary conditions represents limitations that force continuity across the cavity surface both for the potential and its gradient. Several methods are useful to solve the Poisson problem such as the *Apparent Surface Charge* (ASC), *Multipole Expansion* (MPE) and *Generalized Born Approximation* (GBA). In this thesis all the solvation treatments use an ASC approach. As commonly done in the continuum methods, the potential V is separated into two contribution

$$V(X) = V_M(x) + V_\sigma(x) = \int_{R^3} \frac{\rho_M(y)}{|x-y|} dy + \int_\Sigma \frac{\sigma(S)}{|x-s|} ds \tag{2.12}$$

The ASP approach is oriented to solve the apparent charge problem computing $\sigma(s)$ through the discretization in *tessere* of the cavity surface and centering on each of them a weighted and punctual charge.

One of the contribution in which we are particular interested is the G_{rep}. This represents the free energy due to the exchange contribution appearing when QM level of theory is used to describe the electronic state. Although this contribution should be computed by discrete repulsive interaction between the solute-solvent electrons, in the PCM (and in all the solvation continuum models) it is evaluated by the $\hat{Q}_{rep}$ operator based on the solvent density, the numeral density of electron pairs in the solvent, the perpendicular component to the cavity of the electric field generated by the solute and the overlap function. This operator is one electron character and it is used like a discrete surface integral evaluated on each *tesserae* that cover the cavity surface. All the argumentations that support the goodness of the $\hat{Q}_{rep}$ are detailed in Ref. (46). This contribution to the total free energy will be under study in the next chapters and in particular in the Section 3.1 where the repulsion energy will become the most important and dominant contribution due to extremely high pressure. This last implies drastic increasing in the numeral density both for solute and solvent electrons,nevertheless the exponential growth of the electron overlap.

The dynamical polarizability is the basis of the third contribution to the free energy in solution namely G_{dis}. The dispersion energy is described by the $\hat{Q}_{dis}(\overrightarrow{r}, \overrightarrow{r'})$ which is a one- and -two electrons operator like. Both the two terms summed to give $\hat{Q}_{dis}(\overrightarrow{r}, \overrightarrow{r'})$ depend on the average electronic distribution of the solvent. Several parameters are needed to compute this term, such as the solvent refractive index, the first ionization potential and the average transition energy of the solute.

Excluding the cavity formation energy, the other terms are all evaluated as sum of discretized contribution on the cavity surface and all of them are computed in a representative position namely *centroid*. It is very important to understand which kind of results we obtain by a PCM calculation in particular if we want to compare this value with experimental data. When we add the thermal motion term G_{tm} to the others, we obtain an absolute free energy of the solute in that solvent. This value is related to a reference state which is represented by the unperturbed solvent and the among of non-interacting electrons and nuclei of the solute. An estimation of the solvation free energy ΔG_{sol} of the molecule M is given by difference between the PCM free energy computed value and the same quantistic level energy computed *in vacuo* of M. It is important to point out that this estimation is not in principle correct and sometimes not good enough to compare with experimental values. As we will show in Section 3.3 when a very high accuracy below 0.5 Kcal/mol is required, a slightly different scheme must be take into account.

One of the most critical issue in the application of PCM, is the implementation of analytical derivatives of the PCM free energy with respect to cartesian coordinates. The analytical derivatives theory is one of the most important achievement in computational chemistry providing a powerful tool to calculate several molecular properties and the forces acting on the nuclei needful to minimize the molecular geometry. Although exact analytical expressions of derivatives in the gas phase are available, it is challenging to have one in the implicit solvent framework.

The real problem in computing the derivatives of the total energy is that in this term a non analytical computed contribution is included. In fact the reaction field contribution $E_R(\rho, \rho')$ is evaluated numerically hence approximately with $E_R^{app}(\rho, \rho')$. If this term depends on n real parameters it is possible to evaluate partial derivatives respect to some of these in order to obtain important molecular properties. For example if these parameters are the cartesian coordinates the partial derivatives of the total energy will represent the forces acting on each nucleus and hence the gradient needful to evaluate the geometry optimization. There are of course two principal different ways to compute the first derivative of the reaction field energy:

1. Develop the analytical expressions for the exact reaction field energy term and than compute an approximate value for the $(\partial/\partial\lambda_i)E_R(\rho, \rho')$

2. Compute the analytical derivative of the approximate reaction field E^{app}.

One of the most important weak point in the geometry optimization in the solvation continuum models is the non continuity of the first derivative of the energy with respect to the cartesian parameters. Cause of this, very active research in the field of geometry optimization in solution, is still in progress.

So far, we have introduced the principal concepts of the energy contributions in the PCM and the most important meaning of compute a derivative of the reaction field energy. Now we are going to describe briefly the cavitation models and some technical flavors related to the building up of the area which is going to host the solute molecule. This paragraph is indeed a dramatically important argument because it is the basis of developments in the Section 3.3 where the modeling of the cavity shape and size is primary for describe with high accuracy the nuclear bases acidity.

In the solvation continuum models the molecular cavity is the portion of space within the surrounding medium that is occupied by the solute molecule. The molecular surface is indeed the boundary between these two areas. The easiest way to build the cavity is to use a sphere or an ellipsoid with radius or axes dimension described like parameter or parameters. This very simple approach is still in use cause of the availability of exact analytical solution of electrostatic equations. The most common way to define molecular cavities is to interlock a set of spheres each centered on the atoms of the solute. Once molecular cavity is built there are several way to define the molecular surface:

1. The *van der Waals surface* (VWS) is a molecular surface obtained by interlocked spheres centered on each atom and having as radius the corresponding van der Waals radius. This definition is particularly useful when the cavitation energy contribution has to be studied.

2. The *solvent-accessible surface* (SAS) (47) is an extension of the VWS, in fact it is defined as the surface identified by a rolling spherical solvent probe on the van der Waals surface. The probe dimension depends on the molecular solvent properties and nature. This method is usually utilized when short range interactions have to be computed.

3. The *solvent-excluded surface* (SES) (48) has similar character of the SAS procedure but it defines a surface where only the contact point between the cavity and the probe are included. So, this method, defines a smoothed surface across which no solvent molecule can move in. The volume enclosed in the SES but not in the VWS is called *solvent-excluded volume*

In the past years several methods has been developed to approximate the solvent-excluded volume (e.g. MSDOT, DefPol, BLMOL under BLSURF), but the most used in quantum mechanic application is the GEPOL firstly elaborated in Pisa by Tomasi and Pascual-Ahuir (49). In this approach the excluded volume is defined by a set of added spheres putted on the cavity with a recursive algorithm. That procedure include recursively new sets of spheres and it is in principle infinite as it tries to fill a concave space with convex objects. However a threshold level on the added sphere radius is defined and a maximum overlap between two new spheres also force the recursive procedure to stop.

As we mentioned before, the reaction field generated by polarization and all the others non electrostatic contribution to the energy, are computed starting from the discretization of the molecular cavity (surface). This procedure is the *tassellation* and it has the purpose of divide in several subsets of surfaces named *tesserae* with a surface area a, a sampling point $\overrightarrow{s}$ and a unit outward vector $\hat{n}$ at the sampling point. These elements are the main quantities used to compute numerically the PCM energy contribution and derivatives with respect related parameters.

2.3.2 The QM Formulation of the Polarizable Continuum Model

As we told in the previous paragraph, the operator which describe a molecular system in solution in the solvation continuum models framework is an effective Hamiltonian. The solution of this problem refers to the solution of the Schrodinger equation of this last operator. As usual in electronic structure problems we adopt the Born-Oppenheimer approximation which allow the partition of the Hamiltonian into an electronic and a nuclear part as well as the factorization of the solute wavefunction. The Schrodinger equation for the effective Hamiltonian $\hat{H}_{eff}$ is

$$\hat{H}_{eff}|\Psi\rangle = E|\Psi\rangle \tag{2.13}$$

where $\hat{H}_{eff}$ is the sum of the Hamiltonian of the molecular part of the solvation continuum model (M) and the solute-solvent interaction:

$$\hat{H}_{eff} = \hat{H}^0_M + \hat{V}_{int} \tag{2.14}$$

Several way and methods are available to define $\hat{V}_{int}$, however in this thesis we only deal in the ASP corresponding formulation. The operator $\hat{V}_{int}$ can be divided into the four principal components that generate the interaction field, having two-, one- and zero-electron structure. The reaction potential has an electronic and a nuclear source, respectively due to ρ^e_M and ρ^n_M. Once defined the reaction potential and the charge distribution, also the corresponding interaction energy is determinate:

$$U_{int} = U^{ee} + U^{en} + U^{ne} + U^{nn} \tag{2.15}$$

where U^{xy} is the interaction energy between x and y sources. Finally the structure of the effective Hamiltonian can be write more explicitly denoting that with the superscript R the related operator is connected to the solvent reaction potential, whereas with the subscripts r and rr' the one- and two-electron operator:

$$\hat{H}_{eff}|\Psi\rangle = [\hat{H}^0_M + \hat{\rho}^e_r \hat{V}^R_r + \hat{\rho}^e_r \hat{V}^R_{rr'} \langle\Psi|\hat{\rho}^e_{r'}|\Psi\rangle]|\Psi\rangle = E|\Psi\rangle \tag{2.16}$$

It is important to note that the term $\hat{\rho}^e_r \hat{V}^R_{rr'} \langle\Psi|\hat{\rho}^e_{r'}|\Psi\rangle$ introduces a nonlinear character to the equation 2.16 and it is related to the U^{ee} contribution energy. If we follow a canonical order to get molecular wave functions we will define the related HF operator for the solvation continuum model problem. In particular we expand this operator over a finite basis set χ and the Fock matrix reads:

$$\mathbf{F} = \mathbf{h}^0 + \mathbf{G}^0(P) + \mathbf{h}^R + \mathbf{X}^R(P) \tag{2.17}$$

Comparing the equations 2.16 and 2.17 the $\mathbf{h}^0 + \mathbf{G}^0(P)$ is the $\hat{H}^0_M$ term whereas $\mathbf{h}^R + \mathbf{X}^R(P)$ represents $\hat{\rho}^e_r \hat{V}^R_r + \hat{\rho}^e_r \hat{V}^R_{rr'} \langle\Psi|\hat{\rho}^e_{r'}|\Psi\rangle$. For a more detailed description of the HF approach and the QM formulation we remind to Ref. (50).

Finally we want to define the basic energy quantity, namely the *free energy functional*. As for the isolated molecule, also for the solute the Fock operator defined in equation 2.17 is obtained by minimizing an appropriate functional. This last, expressed in a matrix form, reads:

$$G^S_{HF} = tr\mathbf{P}\mathbf{h}^0 + \frac{1}{2} tr\mathbf{P}\mathbf{G}^0(P) + \frac{1}{2} tr\mathbf{P}(j + y) + \frac{1}{2} tr\mathbf{P}\mathbf{X}(P) + V_{nn} + \frac{1}{2} U_{nn} \tag{2.18}$$

The wave functions that make the free energy functional 2.18 stationary are also solutions of the effective Schrodinger equation 2.16. The effective Hamiltonian explicit depends on the charge distribution of the solute expressed in terms of ρ_M^e and thus it is nonlinear. This is the origin of a different free energy functional structure for the HF problem in solution if compared with the analogous in the gas phase. As we will principally study chemical system at the DFT level, it is important to point out that this level of theory does not require modifications in the solvation terms with respect to the HF formalism presented above. DFT is also tractable like HF approach concerning the analytical derivatives of the free energy functional and continuum solvation models coupled to DFT are becoming the routine approach for studies of solvated systems.

3

Developments and Applications

3.1 Towards the elaboration of a QM method to describe molecular systems under very high pressure

3.1.1 Introduction

Continuum solvation methods have gained a wide popularity in the description of solvent effects on the electronic structure, geometry and properties of molecules in solution. The continuum approach has also been extended to other media, as polymers and crystals. There are, however, fields for which this approach has not yet been applied. Among them there is solvation at high pressure. In the last years, important progresses have been made in the experimental studies on molecular systems subjected to very high pressures (51; 52; 53; 54; 55; 56; 57; 58; 59). There are abundant reasons for this effort which requires the availability of specialized structures, but the progresses show that the efforts are rewarding. A few ab initio CarrParrinello simulations of systems at high pressure have appeared (51; 54), with interesting results but at a considerable computational cost. What it lacks is the availability of accurate computational methods of cost comparable with that of isolated molecules, to support, confirm and also predict the outcome of experimental studies.

In this section, we present a preliminary version of a new computational method. The procedure is based on the polarizable continuum model (PCM) (60), calibrated and amply tested for solutions at the standard thermodynamic conditions of pressure and temperature. The present study is focused on two critical items: the definition of the pressure (p) in such extreme conditions, and the elaboration of an analytical code for the calculation of molecular energy gradients for the present applications to the regime of high pressures.

The pressure is a macroscopic thermodynamic property not immediately transferable into QM methods addressing the detailed description of a single molecule. Here, we propose a possible solution for this problem using an analytical formulation of the energy derivatives with respect to nuclear coordinates (notably the first derivatives, giving the gradient).

The most important arguments of this section are organized as follows: the computational method for the calculation of the pressure is presented in Section 3.1.2; the analytical formulation of the molecular energy gradients is described in Section 3.1.3; an application to the study of the effects of a high pressure on the equilibrium geometry and conformational energies of 1,3-butadiene is presented in Section 3.1.4.

3.1.2 Definition of Pressure

The quantum-mechanical description of confined systems has a long history, which goes back to the early days of Quantum Chemistry (61; 62; 63). The pressure for a molecular system confined within a volume V is defined as (64; 65; 66; 67)

$$p = -\left(\frac{\partial E}{\partial V}\right)_{occ.numm.} \tag{3.1}$$

where E is the energy of the system, and the lower script *occ.numm.* denotes a quantum-mechanical adiabatic transformation, i.e., a transformation in which the occupation number of the electrons of the molecular system is kept fixed. The pressure given by Eq. 3.1 is called kinetic pressure, and it has been demonstrated that it is equal to the pressure derived from the quantum statistical mechanics, i.e. the thermodynamics pressure (65; 67; 68; 69; 70).

Up to now the application of Eq. 3.1 has been limited to simple confining volumes, of spherical or ellipsoidal shape only, and for atoms or very simple molecules (71; 72; 73; 74; 75; 76; 77; 78). The introduction of the explicit consideration of pressure given by Eq. 3.1 into PCM permits to remove these limitations. In fact, the volume of the PCM cavity in which the molecule is inserted may be freely reduced, while retaining its shape. In such a way, the following modified form of Eq. 3.1 may be used to monitor the pressure of the external medium acting on the molecule:

$$p = -\left(\frac{\partial G_{el-rep}}{\partial V}\right)_{occ.numm.} \tag{3.2}$$

where G_{el-rep} is the basic energetic quantity of the PCM model (see Eq. 2.3). The derivative of Eq. 3.2 may be computed using the analytical formulation of the volume derivative (79) applied to the analytical derivatives of G_{el-rep} given in the following Section 3.1.3 (more detailed discussion has been developed into my Graduation Degree Thesis titled "Studio di sistemi atomici e molecolari in condizioni di confinamento alle alte pressioni mediante il Modello del Continuo Polarizzabile" University of Parma Italy 2004). With respect to the other simple confining models, the nature of QM Pauli repulsion contribution to the PCM free-energy functional G_{rep} allows confining of the electronic charge distribution of the solute in a smooth way, according to the term used by Le Sar and Herschbach in the concluding remarks of Ref. (73). To compute G_{el-rep} and its derivatives we need to redefine for the appropriate pressure some of the physical parameters used in PCM, namely the dielectric permittivity of the medium and its numeral density, and the size of the molecular cavity. An example of such selection of parameters will be given in Section 3.1.4.

3.1.3 Analytical Gradient

The continuum solvation methods have, as basic and preliminary goal, the determination of the free energy of the system (solute plus solvent). Methods based on a QM description of the focused subsystem (namely the solute) start from here to determine the other properties of the solute, electronic distribution, equilibrium geometry, and molecular observables. We shall make use of the PCM methodology and terminology because this was the first to present a complete dissection of this energy into terms with a well-defined physical meaning, and because it is the method we have introduced and refined in the years. The decomposition of the PCM free energy has been yet reported in Eq. 2.9. In that were reported the electrostatic (and electronic) term contribution, the dispersion term, the repulsion term, the cavity formation term, and the contributions due to thermal motion of the solute (vibrations and rotations), respectively. The justification and the detailed descriptions of these terms are available in the abundant scientific literature, in a number of review papers and in some books. Reference is made here to the most recent review (60) and book (80). For a short but complete description we remind to Section 2.3.

All contributions are influenced, directly or indirectly, by the cavity size and by the medium macroscopic parameters, T and p. As far as the pressure is concerned the key feature is the size of the cavity hosting the solute. We report in this thesis a QM formulation at the self-consistent

field (SCF) level of the G functional expressed in terms of matrices running on a finite basis set $\{\chi\}$ of atomic orbitals used in the expansion

$$G^{SCF}(R) = [tr\mathbf{Ph^0} + \frac{1}{2}tr\mathbf{PG(P)} + tr\mathbf{Ph^{el}} + \frac{1}{2}tr\mathbf{PX^{el}(P)} + V_{nn} + \frac{1}{2}U_{nn}] + tr\mathbf{Ph^{rep}} + tr\mathbf{Ph^{dis}} + \frac{1}{2}tr\mathbf{PX^{dis}(P)} \quad (3.3)$$

where the terms within square brackets correspond to G_{el}, the next one to repulsion, and the last two to dispersion contributions. Lower case matrices $\mathbf{h}^x$ collect one-electron integrals, capital matrices $\mathbf{X}^x$ contain two-electron integrals. The $\mathbf{P}$ matrix describes in this basis the one-electron distribution of the solute. We have not reported in this expression the terms corresponding to the cavity formation and to the thermal motion contributions to the energy because they do not belong in the QM functional to be variation ally minimized. The reader is referred to the already quoted publications for more details.

We just remark a point, because it is pertinent with what will follow. The elements of all the $\mathbf{h}^{el}$ and $\mathbf{X}^{el}\mathbf{(P)}$ matrices are the sums of integrals referring each one to the representative point of a tessera of the cavity surface. In fact, this surface is partitioned into a number K of tesserae with suitable shapes to get the solution of the pertinent integro-differential equations with the aid of the BEM techniques. Number of tesserae and position of the representative points are the same for all the components. The analytical expressions of the derivatives with respect to a given parameter ξ (a Cartesian coordinate of a nucleus in our case) can be obtained via differentiation of the G^{SCF} functional of Eq. 3.3 (81). Available PCM codes only give the analytical expressions of the first (and second) derivatives of G_{el} (the terms within square brackets in Eq. 3.3) (82; 83; 84; 85; 86; 87). We refer again to the quoted papers for more details.

In standard calculations, dispersion and repulsion contributions to the gradient are separately computed starting from semi empirical expressions (88; 89). This mixing of analytical ab initio and semi empirical calculations was dictated by practical considerations, namely the comparison of the computational cost (quite moderate for repulsion, high for dispersion) with the improvement of the desired results. When the desired results are limited to the obtention of energy, geometry and molecular properties for the ground state in standard ambient conditions, semi classical expressions may be safely used, without detriment of the results and at a sizeable

reduced computational cost (90). Things are different in other cases, as the study of pressure effects is considered here.

A quantum version of the repulsion term, inserted in the Hamiltonian is compulsory here (the dispersion term, in contrast, is only indirected affected by pressure). For this reason, we report here the expression of G_{rep} and of its related matrix operator $\mathbf{h}^{rep}$, already implemented in the most recent PCM versions, passing then to the elaboration of its first derivative. The QM Pauli repulsion contribution G_{rep} is given by the functional form (91)

$$G_{rep} = \alpha \int_{r \notin C} \rho_M(r) d^3 r \tag{3.4}$$

with

$$\alpha = \frac{4\pi}{0.7} \rho_S n_{val}^S$$

where the integral in Eq. 3.4 denotes the electronic charge density lying outside the cavity, 0.7 is a numerical parameter (in a.u.) obtained from super-molecule calculations, ρ_S is the numeral density of the solvent and n_{val}^S is the number of its valence electrons. At the SCF level, G_{rep} is obtained as

$$G_{rep} = tr\mathbf{Ph}^{\mathbf{rep}} \tag{3.5}$$

where the elements of $\mathbf{h}^{rep}$ are given by

$$h_{\mu\nu}^{rep} = \alpha(S_{\mu\nu} - S_{\mu\nu}^{in}) \tag{3.6}$$

here $S_{\mu\nu}$ is an element of the overlap matrix over the expansion basis $\{\chi\}$ and $S_{\mu\nu}^{in}$ is given by

$$S_{\mu\nu}^{in} = \frac{1}{4\pi} \sum_k a_k \mathbf{E}_{\mu\nu}(s_k) \cdot \mathbf{n}_k(s_k) \tag{3.7}$$

where the summation runs over the tesserae, and a_k, $\mathbf{E}_{\mu\nu}(s_k)$, $\mathbf{n}_k$ are, respectively, the area of the k-th tessera, a matrix elements of the electronic electric field integrals evaluated at the tessera representative point $\mathbf{s}_k$, and the unit vector normal to the tessera at the same point.

The derivative term of the repulsion operator elements will be indicated by adding the additional superscript ξ to the element. By differentiation of such elements in Eq. 3.6 we have

$$h_{\mu\nu}^{rep,\xi} = \alpha(S_{\mu\nu}^{\xi} - S_{\mu\nu}^{\xi}) \tag{3.8}$$

where $S_{\mu\nu}^{\xi}$ is the derivative of the overlap integral element, while the derivative element $S_{\mu\nu}^{\xi}$ may be expressed in the following way:

$$S_{\mu\nu}^{in,\xi} = \frac{1}{4\pi}\left[\sum_k a_k^{\xi}\mathbf{E}_{\mu\nu}(s_k)\cdot\mathbf{n}_k(s_k) + a_k\mathbf{E}_{\mu\nu}^{\xi}(s_k)\cdot\mathbf{n}_k(s_k) + a_k\mathbf{E}_{\mu\nu}(s_k)\cdot\mathbf{n}_k^{\xi}(s_k)\right] \tag{3.9}$$

where $\mathbf{E}_{\mu\nu}^{\xi}(s_k)$ is the derivative of electric field AO integrals, and $\mathbf{a}_k^{\xi}$ and $\mathbf{n}_k^{\xi}$ are, respectively, the derivatives of the area of the k-th tessera and of pertinent normal vector.

We also present a semi-analytical formulation of the problem, which follows more closely the original formulation of the Pauli repulsion theory given in Ref. (91). In such a formulation, the matrix $S_{\mu\nu}^{in}$ elements of Eq. 3.7 are approximated by a finite difference expression

$$S_{\mu\nu}^{in,\xi} = \frac{a_k}{\delta}\left[V_{\mu\nu}(s_k) - V_{\mu\nu}(s_k + \delta\mathbf{n}_k)\right] \tag{3.10}$$

where $V_{\mu\nu}(s_k + \delta\mathbf{n}_k)$ is the electrostatic potential integrals computed at the point $s_k + \delta\mathbf{n}_k$ (with $\delta = 1x10^{-5}$). By extension of this formulation to the last two terms of the derivative integrals $V_{\mu\nu}(s_k + \delta\mathbf{n}_k)$ in Eq. 3.9, we arrive at the following expression:

$$a_k\mathbf{E}_{\mu\nu}^{\xi}\cdot\mathbf{n}_k(s_k) + a_k\mathbf{E}_{\mu\nu}(s_k)\cdot\mathbf{n}_k^{\xi}(s_k) = \frac{a_k}{\delta}\left[V_{\mu\nu}^{\xi}(s_k) - V_{\mu\nu}^{\xi}(s_k + \delta\mathbf{n}_k)\right] \tag{3.11}$$

where $V_{\mu\nu}^{\xi}(s_k)$ and $V_{\mu\nu}^{\xi}(s_k + \delta\mathbf{n}_k)$ are the derivatives of the electrostatic potential integrals. The explicit expression for the derivative integral $V_{\mu\nu}^{\xi}(s_k + \delta\mathbf{n}_k)$ is given by

$$V^{\xi}_{\mu\nu}(s_k + \delta\mathbf{n}_k) = -\int [\chi^*_\mu(\mathbf{r})\chi_\nu(\mathbf{r})]^{\xi} \frac{1}{\mid \mathbf{r} - (s_k + \delta\mathbf{n}_k)\mid} dr^3$$
$$-(s^{\xi}_k + \delta\mathbf{n}^{\xi}_k)\left[-\int \chi^*_\mu(\mathbf{r})\chi_\nu(\mathbf{r}) \frac{\mathbf{r} - (s_k + \delta\mathbf{n}_k)}{\mid \mathbf{r} - (s_k + \delta\mathbf{n}_k)\mid^3} dr^3\right] \tag{3.12}$$

and a similar expression with $\delta = 0$ holds for $V^{\xi}_{\mu\nu}$. Numerical checks show that the two formulations (i.e. analytical and semi-analytical) agree with each other.

3.1.3.1 Validation of the analytical gradient

Analytical and numerical gradients of the free-energy functional G_{el-rep} have been computed for a suitable set of molecules (Fig. 3.1) at HF level with 6-31g** basis set. The choice of the molecular systems ranges from simple diatomic to more larger systems to explore a statistically significative sample of the accuracy and of the efficiency of analytical gradient algorithm. All the calculations have been performed by using a "water" as solvent (i.e. the value of the dielectric constant of the medium is ϵ=78.), and with the molecular cavities been defined in terms of spheres centered on the atomic nuclei and with radii 1.2 times the corresponding van der Waals radius in Å($r_H = 1.2$, $r_C = 1.7$, $r_O = 1.45$, $r_F = 1.35$). Numerical gradients of the free energy functional have been computed using a numerical steps of 10^{-3} Å. The molecular geometries for the gradient calculations are reported in appendix A of Ref. (92).

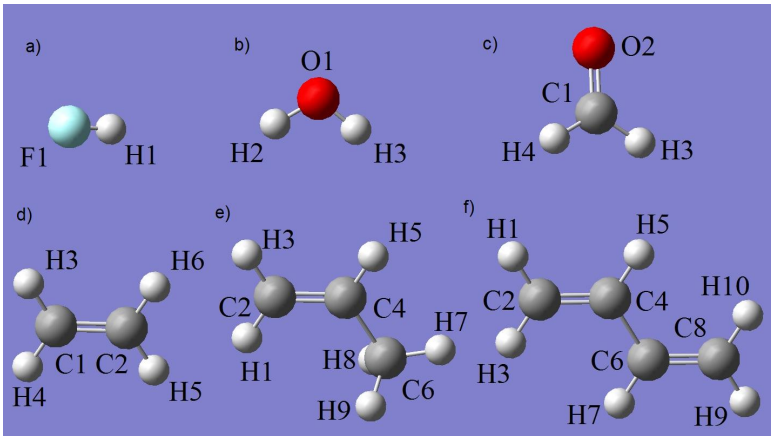


Figure 3.1: Molecular systems under test - Atomic labels for the molecules used in tests:a)HF, b)H_2O, c)H_2CO, d)C_2H_4, e)C_3H_6, f)C_4H_6

3.1.4 Application: high pressure effects on 1,3-butadiene

The analytical gradients for G_{el-rep} have been implemented into a local version of the Gaussian 03 (93) computational code, and validated by comparison with a finite difference method. As an initial application, the computational algorithm has been applied to the determination of the equilibrium geometry and conformational energies of a molecular system subjected at a very high pressure regime. As a molecular system, we have selected 1,3-butadiene for its relevance in the field of high pressure chemistry (54). As already said in Section 3.1.2, to compute the pressure using the derivative of G_{el-rep}, i.e. from Eq. 3.2, we need to redefine some of the physical parameters used in PCM as described below.

3.1.5 Computational Methods

Calculations have been performed for the s-trans, s-cis and skew conformers of 1,3-butadiene. Geometry optimizations in both gas phase and the medium were performed at the DFT/B3LYP level (94) using the 6-31G(d,p) basis set (95). The molecular cavities have been defined in terms of interlocking spheres centered on the atom's nuclei and with radii equal to the corresponding vdW radius (96) multiplied by a scaling factor, *f*. The factor f has been varied within the range (1.4−0.72), being $f = 1.2$ the value used in standard PCM calculations (60). The reduction of f up to 0.7 leads to a gradual shrink of the molecular cavity, with a corresponding increase of the Pauli repulsion G_{rep}. In fact, G_{rep} is proportional to the extent of the electronic charge lying outside the cavity (see Eq. 3.5), which increases when the molecular cavity volume is shrunk.

The electrostatic solute−solvent interaction has been computed using the integral equation formalism formulation of the PCM (IEFPCM) (97). The IEFPCM is in fact able to represent also the effect of the volume polarization density produced by electronic charge lying outside of the cavity, as described by Chipman (98) and Cances and Mennucci (99). The dielectric medium permittivity has been chosen at the constant value of $\epsilon = 2.0$, corresponding to that of liquid butadiene, and the density of the medium has been varied between 0.6 and 1.9 g/ml, on the basis of a recent Carr−Parrinello simulation of compressed butadiene (54).

3.2 Applications of Computational Chemistry in Transition Metal Systems

3.2.1 Luminescence properties of phosphine silver(I) complexes with thiourea derivatives

3.2.1.1 Introduction

The reaction of inorganic silver(I) salts with ternary oregano phosphorus derivatives leads to the formation of different adducts where the metalligand ratio varies from one to four depending on the nature of the ligand, the stoichiometric ratio between the reagents and the experimental conditions (100). These compounds can be usefully employed as precursors of d^{10} metal complexes with mixed ligands having different stoichiometries and geometries. These complexes often show interesting luminescence properties involving charge transfer excited states, whose nature can be assessed only on the basis of reliable calculations (101). The coordination chemistry of the d^{10} metal centers varies markedly as a function of the electronic and satiric factors peculiar to the ligands. The structural characterization of these compounds is of extreme importance in order to shed light on the factors that influence the reaction mechanism and to gain control over the synthesis towards the formation of complexes with predictable geometric requirements. Moreover, it constitutes the starting point for reliable quantomechanical calculations on these systems.

Recently, our colleagues, reported (ref. (102)) the synthesis and the characterization of derivatives of $[(PPh_3)_2AgNO_3]$ with ligands obtained by condensation of phenylisothiocyanate with 2- aminopyridine and 2-aminomethylpyridine namely 1-phenyl-3-(2-pyridyl)-2-thiourea (Phpytu, **1**) and 1-phenyl-3-(2-methylpyridyl)-2-thiourea (Phmepytu, **2**) (Fig. 3.2).

Most important spectral transition of these two complex regards firsts ten excited states so we limited our attention on them. To understand some luminescence properties of these complexes we performed a characterization of firsts excited states involved in UV/vis absorption and luminescence processes. We made theoretical calculations on the two ligands mentioned before and relative complexes $[Ag(PPh_3)_2(phpty)]^+$ and $[Ag(PPh_3)_2(phmepty)]^+$ (Fig. 3.3). TDDFT approach has been used for calculation with three parameter hybrid functional method (B3LYP)(94) and Los Alamos ECP plus DZ (LanL2DZ) basis set for all calculation using the *Gaussian 03* package(93). We decided to start our calculation beginning from experimental X-Ray structure both for ligands and complexes followed by standard geometry optimization

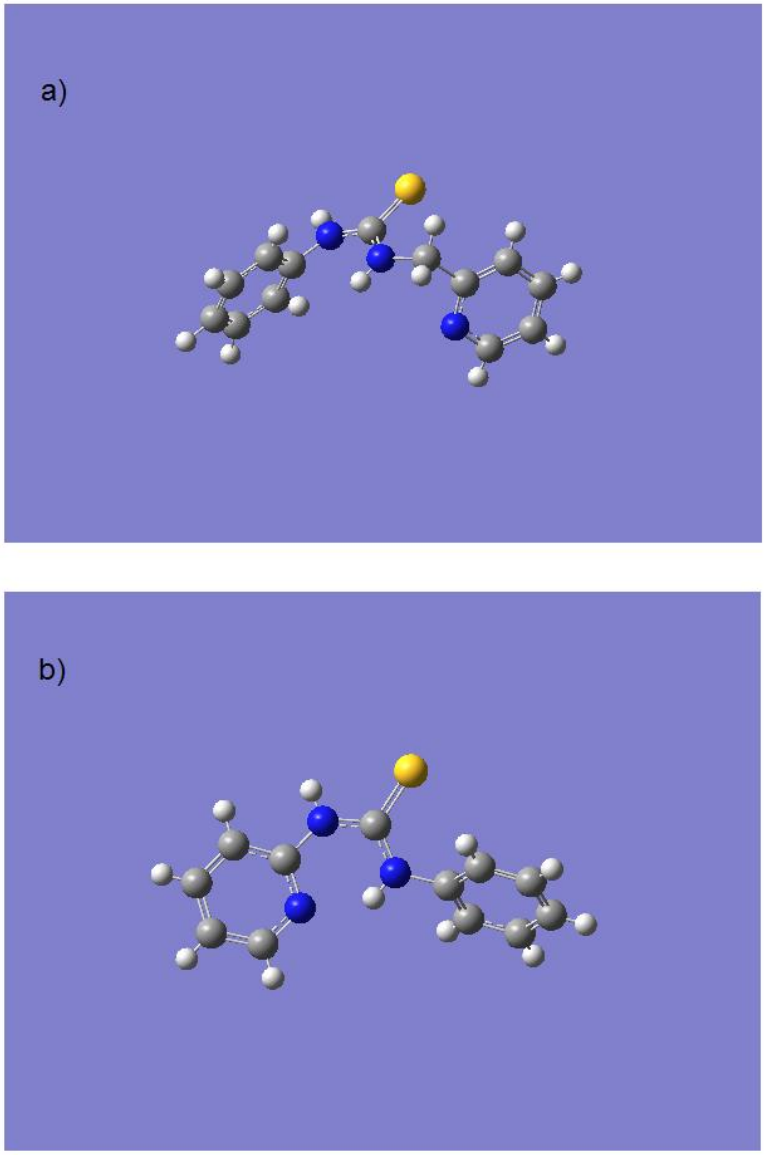


Figure 3.2: Phmepytu and Phpytu ligands - Structure of a) Phmepytu and b) Phpytu

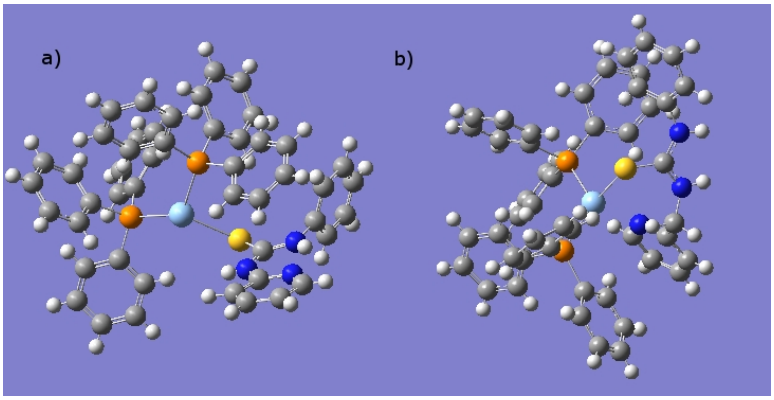


Figure 3.3: Phpytu and Phmepytu silver complexes - Structure of a)Phpytu and b)Phmepytu silver(I) complex

at the computational level described above. In the results chapter 4.2.1 related to this section we will see that the energy and the composition of MOs can give useful characterization of excited states and thus photon induced transitions.

3.2.2 Mechanistic Insights into Acetophenone Transfer Hydrogenation Catalyzed by Half-Sandwich Ruthenium (II) Complexes Containing: a Computational Study
2 - (Diphenylphosphanyl)- aniline

3.2.2.1 Introduction

Transition-metal complexes containing hybrid amino phosphane ligands, generally indicated as P,N ligands, are among the most studied homogeneous organometallic catalysts (103; 104; 105; 106; 107; 108; 109). Their success is often attributed to the combination of the soft (P) and hard (N) character of the two donors, which assures a good balance between robustness and reactivity. A well known P,N ligand is 2-(diphenylphosphanyl)- aniline (hereafter referred as PNH_2), whose coordination capability towards several transition-metal ions is known (110; 111; 112; 113; 114). Usually, PNH_2 chelate the metal in a κ^2P,N way (111; 112; 115; 116; 117; 118; 119), although κ^1P(114) and bridging(110; 114) coordination modes are also reported.

Deprotonation of the amine functionality occurs in the presence of a base, giving rise to amido phosphane complexes, where the anionic ligand chelates in a κ^2- P,N way (115; 118). Half-sandwich ruthenium(II) complexes containing amino phosphane ligands draw the interest of several groups (120; 121; 122; 123), especially for their potential catalytic applications (123;

124; 125; 126; 127; 128; 129; 130; 131). Among these, the hydrogen transfer reaction (HTR) of ketones represents an important example (132). Recently, Ikariya reported on reduction processes(127; 129) and racemization of nonracemic sec-alcohols(125) catalyzed by neutral complexes of the type [Cp*Ru(κ^2P,N)Cl], where P,N are several (primary) amino phosphane ligands, PNH_2 included. A metalligand bifunctional catalysis(133; 134) has been proposed on the basis of the observation that similar complexes containing (dimethyl)amino phosphane ligands turned completely inactive in the same processes. The presence of NH_2 functionalities is not, however, strictly necessary to construct active pre catalysts.

Neutral and ionic complexes of the type [(arene)Ru(κ^2P,N)X]$^{n+}$ (arene = Cp or pcymene; P,N = several N,N-dimethylaminophosphanes; X = CH_3CN, Cl, or Br; n = 0, 1) have led from moderate to good activities in the HTR of several ketonic substrates (124). Moreover, the Turn over Frequencies (TOF) values up to 220000 h^{-1} were reached with the zwitterionic complex [Ru(p-cymene)(κ^2P,N)Cl] (131), where the ligand 1-diisopropylphosphanyl-2-(N,N-dimethylamino)- 1H-indene is void of NH functionalities. In these cases, an inner-sphere mechanism (ISM) (135; 136; 137; 138; 139) seems more likely.

On the contrary, Baratta recently demonstrated that pre catalysts containing NH_2 functionalities do not automatically operate through an outer-sphere mechanism (OSM) in HTR (140; 141). Kinetic data and isolation of isopropoxide species suggest in fact that the ruthenium(II) complex [RuCl(CNN)(dppb)] [dppb = $Ph_2P(CH_2)_4PPh_2$; HCNN = 6-(4'-methylphenyl)-2-pyridylmethylamine] works essentially through an ISM, where the NH_2 functionality is, however, involved in the activation of the ketone through N$-$ H$\cdots$ O=C intermolecular hydrogen bonds.

To the best of our knowledge, a detailed mechanistic study on the HTR of ketones catalyzed by half-sandwich ruthenium(II) complexes bearing P,N ligands has not yet been reported. For this reason, here we report on the computational characterization of a ruthenium(II) complex of the type Ru[(κ^2P,N)PNR_2](p-cymene)Cl (R = H) that have been tested as homogeneous pre catalysts for the HTR of acetophenone. A detailed DFT-PCM mechanistic study based on MS (ESI) and spectroscopic data was carried out in order to elucidate the role played by the NH_2 functionality in the catalytic process.

HTR of ketones catalyzed by half-sandwich ruthenium- (II) complexes in alcoholic media has been deeply investigated by theoretical calculations in order to clarify the intimate steps through which the alcohol and ketone interact with the metal center. Most of these works have been performed on model catalysts containing diamine or amino alcohol ligands

(133; 142; 143; 144; 145; 146), whereas to the best of our knowledge, no reports dealing with aminephosphane based catalysts have so far been reported. Initially we undertook a computational investigation on the conformational stability of **1** in vacuo as well as in iPrOH solution to establish the factors governing the ring-closure process observed in solution that transforms **1** into **1a** (Fig. 4.7). Then, the study was directed to the OSM promoted by **1a**, which is hereafter referred to as Path I. In order to have a more complete picture, we also modeled the other two mechanisms deriving from the reactions that iPrO might trigger off by reaction with 1a, that is, the formation of the isopropoxide intermediate Ru[(κ^2P,N)- PNH_2](p-cymene)(iPrO)$^+$ (Path II) and the neutral complex Ru[(κ^2P,N)PNH](p-cymene)Cl (Path III). These two last species would promote the HTR by an ISM.

Finally, we compared the corresponding energy profiles to establish the most energetically favored mechanism. In order to have an accurate picture from calculations, the organometallic fragments were built without any simplification, and iPrOH and acetophenone were used as a hydrogen donor and acceptor, respectively. The energy profiles were determined by calculating the ΔG values as a sum of the electrostatic and non electrostatic contributions (60) and a correction for the zero-point energy (147) (for the evolution of the reacting systems see the Supporting Information 8).

3.2.2.2 Computational Methods

The geometries of all minimums and saddle points were optimized in the gas phase at density functional theory level by means of hybrid B3LYP functional (94). For geometry optimization D95V (H, C, N, O) and ECP plus DZ (P, Cl, Ru) basis sets were used (BS-0). Thermo chemical analysis was performed on all intermediates and unique imaginary frequencies searched in transition states. Accurate energy calculation was performed on gas-phase optimized geometry at B3LYP level in solvent environment simulation by using integral formalism polarizable continuum model (IEF-PCM) at five different basis set levels: (a) BS-0; (b) 6-31G (H, C, N, O) and ECP plus DZ (P, Cl, Ru) (BSI); (c) 6-31G(d,p) (H, C, N, O, P, Cl) and ECP plus DZ (Ru) (BSII); (d) 6-31G(d,p) (H, C, N, O, P, Cl) and ECP plus DZ (Ru) (BSIII); (e) 6-31++G(d,p) (H, C, N, O, P, Cl) and ECP plus DZ (Ru) (BS-IV). Only BS-IV data are reported in the resultes chapter related to this section, because different levels of accuracy did not show important variation in the qualitative and quantitative description of the catalytic cycle. Anyway the basis set influence on the mentioned resultes are reported in Appendix 7.11. The PCM parameters are: Ua0 model for building of cavities, dielectric constant fixed to 18.3, density of the medium

equal to 0.007876 particle/Å^3 and 0.3 Å^2 the areas of tesserae. Additional spheres were added on acid hydrogens. All the calculations were performed with the Gaussian03 package (93).

3.2.3 Oxidative Addition of Iodomethane to Charge-tuned Rh(I) Complexes:a frequency and population analysis of the catalyst

3.2.3.1 Introduction

It is established, by experiment and calculation, that oxidative addition of methyl iodide to square planar rhodium(I) carbonyl complexes proceeds via a nucleophilic mechanism, where the rate determining step shows a $SN2$ transition state (148; 149; 150). Consequently, if the reactivity is not strongly moderated by steric effects, due to bulky substituents or ligand orientation that hinder the space above and below the Rh coordination plane(151), higher reaction rates are found for complexes with high electron density on the Rh center, which behaves as a nucleophile towards the carbon atom of CH_3I.

The value of the IR stretching frequency of coordinated CO is taken as an indicative measure of the electron density on the Rh atom, and a qualitative correlation between the electronic density, the CO frequency and the reaction rate is possible (as a rule of thumb, faster kinetics are observed for complexes with a CO frequency of about 1990 cm^{-1} or lower). However, Haynes et al (152) suggest that similar complexes with sensibly different CO stretching frequency, namely trans-[Rh(CO)(PPh_3)$_2$I] and [Rh(CO)(dppe)I] (dppe = 1,2-bis(diphenylphosphino)ethane; ν CO = 1981 and 2011 cm^{-1} respectively), might be expected to have similar electron density on the Rh center.

In our opinion, this shows that the Rh electron density is not only dependent on the kind of donor atoms in the ligands, but also on their orientation and geometry, and that this density is indeed reflected by the CO stretching frequency. Nonetheless, in different complexes with the same donor atoms and the same coordination geometry, the CO stretching frequency can be used, with confidence, as a measure of the electronic densities.

Recently in the Prof. Cauzzi's laboratories, an extensive study of kinetic and thermodynamic properties of the oxidative addition of iodomethane to the zwitterionic metallate [Rh(EtSNS)(CO)] (**c1**) and its protonated derivatives ([Rh (HEtSNS) (CO)]X (2·X) and [Rh (H2EtSNS) (CO)]X_2 (3·X_2) [X = PF_6^-, OTf, NO_3^-; EtSNS = EtNC (S) Ph_2 P=NPPh_2 C(S)NEt^-]) has been carried out. Compound **c1** is a biprotic base and can be protonated affording cationic species $[Rh(HEtSNS)(CO)]^+$ (**c2**) and $[Rh(H2EtSNS)(CO)]^{2+}$ (**c3**) in which protons bind to the nitrogen atoms of the thioamidyl functions [pK_a in CH_2Cl_2 = 6.5(3) and 4.8(4)](153) (Fig. 3.4).

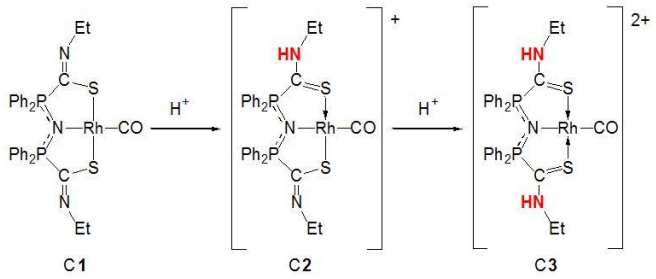


Figure 3.4: Zwitterionic metallate and derivatives - Formation of the mono-cationic complex **c2** and the di-cationic complex **c3** by protonation of the zwitterionic metallate **c1**

In this thesis we are going to study the effects of the protonation of the two possible thioamidyl (N) side on the electron density distribution and hence on the CO frequency stretching. Moreover we will establish a quantitative correlation between CO frequency stretching and electron population on the metal both for isolated complexes and ionic couple with few counter ions.

3.2.3.2 Computational Methods

The geometries of all minimums were optimized in the gas phase at density functional theory (DFT) level by means of hybrid B3LYP functional (94). The geometries were optimized in gas phase with 6-31g** basis set on H, C, N, O, P and S atoms and Lanl2DZ basis and pseudo potential on Rh developed by Hay and Wadt (BS-0). The latter uses a semi core double- contraction scheme for the heavy elements such as Rh. A more wide basis set has been used to evaluate the pseudo potential effects on some molecular properties, substituting the Lanl2DZ functions and potential with complete dgdzvp functions (DGauss Functions) (BS-I)(154; 155). All the simulations were performed both in the gas phase and solution by means of the Polarizable Continuum Model (PCM). The continuum were built in CH_2Cl_2 environment and standard PCM energies were computed with UA0 cavity model. Thermo chemical analysis has been performed on all the minimum and no imaginary frequencies were found. Accurate energy calculation has been performed at the same level with tight SCF criteria on the wavefunction. Mulliken population analysis(156) were performed at each calculation level. The MO rendering was performed with the cubegen utility provided in the Gaussian03 package. All the calculations were performed with Gaussian03 and gdv package (93).

3.3 Calculation of pK_a Values of Nucleobases and the Guanine Oxidation Products Guanidinohydantoin and Spiroiminodihydantoin using Density Functional Theory and a Polarizable Continuum Model

3.3.1 Introduction

Damage to DNA via oxidation is thought to be related to a variety of cancers and neurological disorders as well as cell aging and death (157; 158; 159; 160; 161). Typical damage involves chemical oxidation of the nucleobases or the pentose sugars, base cleavage or the formation of DNA:protein crosslinks. Oxidation of the guanine base leads to a variety of products including 8-oxo-7,8-dihydro-2'-deoxyguanine (8-oxoG), guanidinohydantoin (Gh), and spiroiminodihydantoin (Sp) (162; 163; 164). Gh and Sp are two of the major products observed in guanine oxidation and are thought to be formed via a common intermediate, 5-hydroxy-8-oxo-7,8-dihydroguanosine with the product branching ratio dependent on the pH of the reaction environment (165; 166; 167; 168; 169).

Computational studies conducted by H. B. Schlegel's group (170) confirmed that protonation or deprotonation of various sites within the intermediate species involved in the mechanism of formation of Gh and Sp had a significant effect on the predicted kinetics and thermodynamics of the various pathways. For future studies of acid- or base-catalyzed chemical reactions, it would be useful to have an efficient computational method to predict the site-specific protonation state or pK_a of intermediate species at experimental reaction pHs. In addition to understanding the likely mechanisms for pH sensitive reactions, a theoretical method for computing relative pK_as might prove useful for predicting the relative toxicity or mutagenicity of these adducts.

The structural and functional changes observed in DNA following oxidation of its nucleobases may be due, in part, to changes in the hydrogen-bonding characteristics of the oxidized adducts. Hydrogen bonding between the nucleobases in DNA and RNA duplexes is known to be very important to their structure and function in vivo (171). The strength of these hydrogen bonds is correlated to the relative pK_as of the donor and acceptor nucleobases (172). For several decades, research groups have used a variety of experimental techniques to measure the pK_a of the isolated DNA and RNA nucleobases as well as those attached to the ribose, deoxyribose and to the phosphate backbone (164; 173; 174; 175; 176; 177; 178). The pK_a

of nucleobases has been demonstrated to fluctuate depending on its location in the base sequence in the backbone of DNA and RNA, its proximity to metal ions and on its oxidation state (164; 175; 176).

Research groups have used computational methods to predict either the pK_a or protonation state of both modified (179; 180) and unmodified nucleobases (181; 182; 183; 184; 185; 186; 187) ,nucleotides (188) and the guanine oxidation product, 8-oxoguanine (189). The latter compound can be further oxidized in vivo to form guanidinohydantoin and spiroiminodihydantoin, two highly mutagenic DNA lesions (190; 191; 192; 193). In UV melting studies, both guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp) were shown to decrease the thermodynamic stability of duplex DNA relative to guanine and 8-oxoguanine with the effect of the Sp being more severe (190; 194).

Molecular dynamics simulations indicate that Sp lesions alter the base stacking and Watson-Crick hydrogen bonding interactions of the duplexes (195). While steric differences between the planar guanine and 8-oxoguanine and the non-planar Gh and Sp compounds could certainly contribute to these conformational changes, it is also possible that the duplex destabilization is due, at least in part, to changes in the site specific pK_as of these lesions relative to their parent compounds guanine and 8-oxoguanine. The prediction of accurate pK_as using computational techniques has been the subject of study for many years (172; 196; 197; 198; 199; 200; 201; 202; 203; 204; 205; 206; 207; 208; 209; 210; 211; 212; 213; 214; 215; 216; 217; 218).

It is particularly challenging due to the fact that an error of 1.36 kcal/mol in the free energy calculation results in a error of 1 pK_a unit. A variety of computational methods have been used to study the relative acidity and basicity of the nucleobases, nucleosides, and nucleosides. Theoretical studies of the gas phase acidity of the isolated nucleobases were conducted by several groups. Giese and McNaughton studied the site-specific proton affinities of guanine and its heptahydrate at the B3LYP/6-31++G(d,p) level of theory. The proton affinity was calculated to be 229.7 kcal/mol and found to be consistent with the experimental value of 227.311.5 kcal/mol (187). Leszczynski calculated the proton affinities of the five nucleobases and the rare tautomers of guanine, adenine and cytosine at the MP4(SDTQ)/6-31+G(d,p)//MP2/6-31+G(d,p) and MP2/6-311++G(d,p)//MP2/6-31+G(d,p) levels of theory. The calculated values were reported to be within 2.1 % of the experimental values and their data suggested that the rare tautomers of the nucleobases were a significant portion of the gas-phase equilibrium composition (186). Chandra and Zeegers-Huyskens used B3LYP density functional theory and the 6-31++G(d,p) and 6-311++G(d,p) basis sets (182; 183; 184; 185). to calculate the relative

acidity of various tautomers of cytosine, thymine and uracil both alone and complexed with one molecule of water. They reported the relative acidities of the five nucleobases as uracil >thymine >guanine >adenine >cytosine. The solution-phase pK_a values of nucleotides in RNA have been calculated to within 1 to 2 pK_a units using a modified version of the multi-conformational continuum electrostatics (MCCE) program which employs a variation of the Poisson-Boltzmann equation coupled with Monte Carlo treatment of the multiple ionization states (188). This group predicted that significant shifts in the pK_a of the adenine and cytosine nucleotides could be brought about by changes in the 3-D structure of the RNA backbone and may be responsible for its function.

Chatterjee et al. (178) estimated the effect of modification of the pentose-sugar on the site-specific pK_a values for adenosine, guanosine, cytidine, thymidine, and uridine nucleosides using the closed shell Hartree Fock (HF) level of theory with a 6-31G(d,p) basis set. The conductor-like polarizable continuum model (CPCM) was used to calculate the free energy of solvation for each of the modified nucleosides and the results were compared to experimental data. For these calculations, the free energy of solvation for the proton was assumed to be -263.47 kcal/mol. The authors reported a good linear correlation (R=0.98, pK_a(exp) = 0.4690(0.0170)* pK_a(calc)-2.1087(0.3270) between the predicted pK_a values of the nucleosides and the experimental values obtained with their corresponding bis-ethylphosphate nucleotides.

Using the B3LYP density functional theory and the Poisson-Boltzmann continuum-solvation model with modified atomic radii, Goddard's group has calculated both the site specific and overall pK_a values of guanine, cytosine, isoguanine, 9-methylisoguanine, xanthine and 8-oxoguanine (179; 181; 189) By treating the solvation free energy of the proton as a variable ($\Delta G_{sol}(H^+)$= -263.47 kcal/mol), they were able to predict the pK_a of guanine within 0.2 units of experiment and the values for cytosine, isoguanine, xanthine and 9-methylisoguanine within 1 unit of experiment. Goddard also estimated the relative abundance of deprotonated guanine at physiologic pH and discussed the role of the deprotonated species in base-pair mismatching during DNA replication. Using the same method, Hwang and Jang (180; 219) predicted the pK_a values of 9-methylguanine, 9-methyladenine, and 9-methylhypoxanthine to within 0.6 to 1.5 units of the experimental data.

In this thesis, we propose an efficient computational method for calculating the pK_a of DNA and RNA nucleobases and the guanine oxidation products, guanidinohydantoin and spiroiminodihyantoin employing the thermodynamic cycle outlined in Fig. 3.5, Boltzmann weighting of

relevant tautomers, and modification of the UFF solvation cavity. The thermodynamic cycle is a modification of the one proposed by Nasimento et al. (214) for the calculation of absolute pK_a values for carboxylic acids.

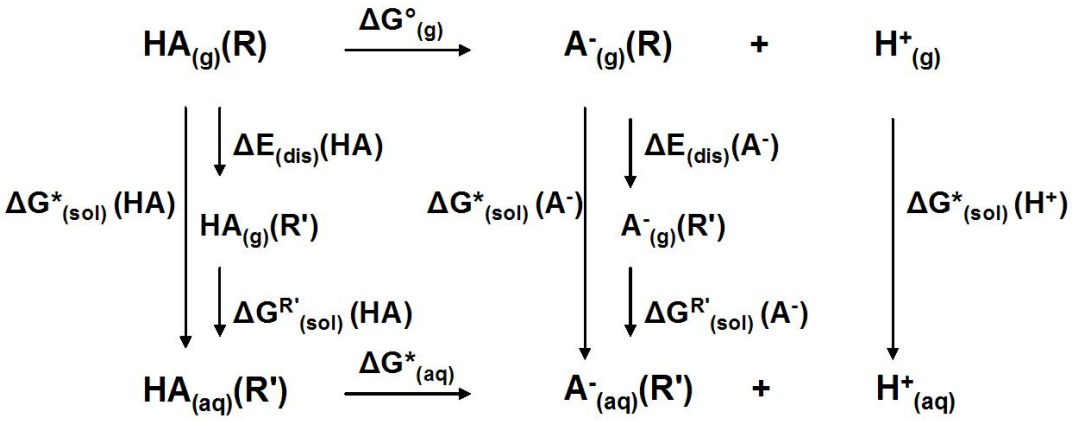


Figure 3.5: Thermodynamic cycle used in the calculation of the pK_a - The standard state reference for the gas phase free energy, $\Delta G^0_{(g)}$, is one atmosphere of pressure and 298.15K and for the free energy in water, $\Delta G^*_{(aq)}$, and free energy of solvation, $\Delta G^*_{(sol)}$, it is 1M and 298.15K. R is the structure optimized in the gas phase at B3LYP/6-31+G(d,p) and R' is the structure optimized in solution at IEF-PCM/B3LYP/6-31+G(d,p). $\Delta E_{(dis)}$ is the electronic distortion energy between the gas and solution phase optimized structures calculated at the same level of theory with the aug-cc-pVTZ basis set. $\Delta G^{R'}_{(sol)}$ is the calculated free energy of solvation of the R' structure

3.3.2 Computational Methods

3.3.2.1 General formulas for calculation of global and site-specific pK_a values

As noted in work previously published by Goddard's group (181; 189), the calculation of pK_a values is complicated by the presence of multiple tautomers having various site-specific pK_as. We have followed the method outlined by Goddard for calculating a global pK_a value from a Boltzmann weighting of the site-specific values of the various tautomers investigated in our study (eq 3.13). Where pK_a^{ij} is the site-specific value for deprotonation of tautomer *i* resulting in the formation of the deprotonated tautomer *j*, and f_i and f'_j are the Boltzmann weighted fractions of tautomers *i* and *j*, respectively.

$$pK_a = pK_a^{ij} - logf_i + logf'_j \tag{3.13}$$

The site-specific pK_a of an acid HA is given by Eq. 3.14 where R is the gas constant, T is the temperature,

$$pK_a = \frac{1}{2.303RT}\Delta G^0_{(aq)} \tag{3.14}$$

and ΔG*(aq) is the free energy of the deprotonation reaction, HA → H^+ + A^-, for a standard state of 1 mol/L and room temperature. Using the thermocycle defined in 3.5, ΔG*(aq) is defined as the difference in the free energies in solution between the proton (H^+) and unprotonated tautomer (A^-) and the protonated tautomer (HA) (Eq. 3.15).

$$\Delta G^0_{(aq)} = G^0_{(aq)}(A^-_{R'}) + G^0_{(aq)}(H^+) - G^0_{(aq)}(HA_{R'}) \tag{3.15}$$

For each species, G*(aq) is the sum of the gas phase standard free energy, G*(g) and the free energy of solvation in water, ΔG*(sol), where the "*" indicates that all terms are in the standard state of 1-mol/L. To convert the calculated gas phase standard free energy, G(g), from its standard state of 1-atm gas phase/1-m solution to G*(g) with a standard state of 1-M gas/1-M solution phase, it is necessary to add 1.89 kcal/mol (RTln(1/R_gT) (220), (Eq. 3.16)

$$\begin{aligned} G^0_{(aq)} &= G^0_{(g)} + \Delta G^0_{(sol)} \\ &= (G^0_{(g)} + 1.89) + \Delta G^0_{(sol)} \end{aligned} \tag{3.16}$$

where R is 1.987 cal/mol · K and R_g is 8.206 x 10^{-2} liter· atm/mol·K. Combining equations 3.14, 3.15 and 3.16 yields a general expression for the calculation of the site-specific pK_a of a given tautomer (Eqs. 3.17 to 3.19).

$$\begin{aligned} pK_a = \frac{1}{2.303RT} \quad & (G^0_{(g)}(A^-) + 1.89 + \Delta G^0_{(sol)}(A^-) + G^0_{(g)}(H^+) + 1.89 + \Delta G^0_{(sol)}(H^+) \\ - \quad & G^0_{(g)}(HA) - 1.89 - \Delta G^0_{(sol)}(HA)) \end{aligned} \tag{3.17}$$

$$\begin{aligned} = \frac{1}{2.303RT} \quad & (G^0_{(g)}(A^-) + 1.89 + \Delta G^0_{(sol)}(A^-) + G^0_{(g)}(H^+) + \Delta G^0_{(sol)}(H^+) \\ - \quad & G^0_{(g)}(HA) - \Delta G^0_{(sol)}(HA)) \end{aligned} \tag{3.18}$$

$$= \frac{1}{2.303RT}(G^0_{(g)}(A^-) + \Delta G^0_{(sol)}(A^-) - G^0_{(g)}(HA) - \Delta G^0_{(sol)}(HA) - 270.29) \tag{3.19}$$

The gas and solution phase free energies of the proton were taken from the literature to be G(g)=-6.28 kcal/mol (196) and ΔG*(sol)=-265.9 kcal/mol (220), respectively.

3.3.2.2 Calculation of gas and solution phase free energies

Molecular orbital calculations were carried out using the E05 development version of the GAUSSIAN series of programs (Note: the PCM parameters change significantly between versions E05 and F01) (221). Optimized geometries and energies in the gas phase and in aqueous solution were computed with the B3LYP density functional method (222; 223; 224) using the 6-31+G(d,p) basis set (224; 225; 226; 227; 228; 229) (BS-I) with the latter calculations also employing integral equation formalism of the polarizable continuum model (IEF-PCM)

(97; 230; 231). Tight convergence criteria and the "nosymm" options were used for all optimizations.

The solution phase optimizations employed a solvent excluding surface cavity model (232), UFF radii (233), and tesserae with an average area of 0.200 Å^2. For this study, all Gh and Sp chiral tautomers were the "R" stereoisomer. Given the symmetric nature of the PCM solvation model, it is anticipated that the calculated pK_a values would be the same for the "S" isomer. Single point calculations (in the gas phase) were also conducted with the gas phase and solution phase optimized geometries with the aug-cc-pVTZ (234) basis set (BS-II) and tight convergence criteria. Cartesian coordinates for the optimized geometries and electronic energies for all compounds are provided in the Supporting Information of Ref. (235). Vibrational frequencies were computed in the gas phase at the B3LYP level with the 6-31+G(d,p) basis set (BS-I) and were used without scaling since the B3LYP frequencies agree quite well with experimental values for a wide range of second and third period compounds (236). Thermal corrections and free energies were calculated by standard statistical thermodynamic methods (237) using the unscaled B3LYP frequencies and the ideal gas / rigid rotor / harmonic oscillator approximations. To improve the accuracy of the pK_a calculated values versus experimental data, the solution phase single point calculation of the free energy of solvation employed a modified UFF cavity using an α value of 0.91 for the cationic and neutral species and 0.83 for the anionic species. Selection of the optimal values was made by the empirical fitting process described below.

3.3.2.3 Calculation of pK_a

Following optimization of each of the species in the gas phase, the standard gas phase free energy, G(g), for each species was calculated as the sum of the electronic energy at 0 K, the unscaled zero point energy, and the change in the free energy from 0 to 298 K (Eq. 3.20). Here R refers to the gas phase optimized structure.

$$G^0_{(g)}(R) = E^{B3LYP/aug-cc-pVTZ}_{0K}(R) + ZPE^{B3LYP/6-31G(d,p)}(R) + \Delta G^{B3LYP/6-31G(d,p)}_{0->298K}(R) \quad (3.20)$$

The standard free energy of solvation in water ΔG*(sol) as described by Ben-Naim et al. (238), is defined as the difference between the gas phase free energy at the gas phase optimized geometry, R, and the solution phase free energy at the solution phase optimized geometry, R' (Eq. 3.21 and 3.22).

$$\Delta G^{0}_{(sol)}(A^{-}) = G^{0}_{(aq)}(A^{-}_{R'}) - G^{0}_{(g)}(A^{-}_{R}) \tag{3.21}$$

$$\Delta G^{0}_{(sol)}(HA) = G^{0}_{(aq)}(HA_{R'}) - G^{0}_{(g)}(HA_{R}) \tag{3.22}$$

In the thermodynamic cycle outlined in Fig. 3.5, the standard free energy of solvation, $\Delta G^{*}_{(sol)}$, is partitioned into two physically meaningful parts (eq. 3.23 and 3.24): a deformation term which captures the electronic distortion energy resulting from the change in the geometry of the solute as the species moves from the gas phase (R) to the solution phase (R'), $\Delta E_{(dis)}$ (eq. 3.25 and 3.26), and a free energy of solvation term for the molecule at its geometry optimized in solution, $\Delta G^{R'}_{(sol)}$ (eq. 3.27 and 3.28). $\Delta G^{R'}_{(sol)}$, was obtained from a single point calculation at the IEF-PCM/B3LYP/6-31+G(d,p) level of theory using the modified UFF cavity. The distortion energy ($\Delta E_{(dis)}$) for each species was computed from single point calculations conducted at the B3LYP/aug-cc-pVTZ level of theory on the gas phase (R) and solution phase optimized (R') structures optimized at B3LYP/6-31+G(d,p) and IEF-PCM/B3LYP/6-31+G(d,p) respectively.

$$\Delta G^{0}_{(sol)}(A^{-}) = \Delta G^{R'}_{(sol)}(A^{-}) + \Delta E_{(dis)}(A^{-}) \tag{3.23}$$

$$\Delta G^{0}_{(sol)}(HA) = \Delta G^{R'}_{(sol)}(HA) + \Delta E_{(dis)}(HA) \tag{3.24}$$

$$\Delta E_{(dis)}(A^{-}) = E^{0}_{(g)}(A^{-}_{R'}) - E^{0}_{(g)}(A^{-}_{R}) \tag{3.25}$$

$$\Delta E_{(dis)}(HA) = E^{0}_{(g)}(HA_{R'}) - E^{0}_{(g)}(HA_{R}) \tag{3.26}$$

$$\Delta G^{R'}_{(sol)}(A^{-}) = G^{0}_{(aq)}(A^{-}_{R'}) - G^{0}_{(g)}(A^{-}_{R'}) \tag{3.27}$$

$$\Delta G^{R'}_{(sol)}(HA) = G^0_{(aq)}(HA_{R'}) - G^0_{(g)}(HA_{R'}) \tag{3.28}$$

This approximation has been employed for computational reasons, as it avoids the calculation, often problematic, of the normal modes in solution. As a check of the validity of this approximation, the partition functions for the gas phase and solution phase optimized geometries of two cationic, two neutral and two anionic species were evaluated. The mean absolute deviation in the calculated relative free energy, ΔΔG, of the species was found to be 0.1 to 0.7 kcal/mol.

Semi-empirical modification of the cavity scaling factor to more accurately predict the experimental pK_a. values should reduce the impact of the differences in the partition functions. When the standard free energy of solvation is expressed in terms of the distortion energy and the solvation energy of the distorted geometry, $\Delta G_{(sol)} = \Delta E_{(dis)} + \Delta G^{R'}_{(sol)}$, equation 3.19 becomes

$$pK_a = \frac{1}{2.303RT}(G^0_{(g)}(A^-_R) + \Delta G^{R'}_{(sol)}(A^-) + \Delta E^0_{(dis)}(A^-) - G^0_{(g)}(HA_R) - \Delta E^0_{(dis)}(HA) - \Delta G^{R'}_{(sol)}(HA) - 270.29) \tag{3.29}$$

This equation is used to calculate the site-specific pK_a for each tautomer.

4

Results and Discussions

4.1 Analysis on system under very high pressure

4.1.1 Analytical gradient validation test

Tables 4.1, 4.2 and 4.3 report the numerical results for the free-energy analytical and numerical gradients G_{el-rep}, with the corresponding computational times for the HF, H_2O, H_2CO, C_2H_4, C_3H_6 and C_4H_6 systems. Let us first consider the accuracy of the analytical and numerical gradients of the free energy, both computed in out from equilibrium geometry for each molecule. Numerical value for HF, H_2O, H_2CO and C_3H_6 show an almost perfect agreement between analytical and numerical gradients. Standard deviation between the two sets of data range from 10^{-5} to 10^{-8} (Hartree/Bohr), and the relative differences are lower then 10^{-2} percent. The others two molecular systems, C_2H_4 and C_4H_6, present a similar pattern of agreement between analytical and numerical gradients with some exceptions, in particular for the component of gradient relative to the C_1 and H_2 atoms in C_2H_4 and C_6 and H_7 for C_4H_6 (for labels see Fig. 3.1. Further tests on the numerical gradients in which we have considered variations of the nuclear step size displacement, and variations of dimension of the cavity and of the degree of the tessellation of its surface, have clearly shown that all the cases of disagreement between analytical and numerical gradients of the free-energy functional are related to the instability of the numerical gradients, which may appear when we include the Pauli repulsion term. These numerical observations clearly point out a more robust behavior of the analytical gradient with respect to its numerical counterpart.

Let us now consider the computational efficiency of the analytical gradients. The computational time required by analytical and numerical gradients has been directly compared in Fig.

HF		Analytical	Numerical	H_2O		Analytical	Numerical
Atom	Coord.	derivatives	derivatives	Atom	Coord.	derivatives	derivatives
	x	0.000000000	0.000000003		x	0.005032221	0.005031057
H_1	y	0.000000000	-0.000000011	O_1	y	0.000076171	0.000076160
	z	0.149649646	0.149649546		z	0.003842874	0.003841444
	x	0.000000000	-0.000000007		x	0.001751204	0.001751198
F_2	y	0.000000000	0.000000004	H_2	y	-0.000009902	-0.000009898
	z	-0.149649646	-0.149649546		z	-0.007577512	-0.007576204
					x	-0.006783425	-0.006782295
				H_3	y	-0.000066269	-0.000066272
					z	0.003734637	0.003734722
	Job time:	7.9	54.2		Job time:	8.4	83.7
H_2CO		Analytical	Numerical			Analytical	Numerical
Atom	Coord.	derivatives	derivatives	Atom	Coord.	derivatives	derivatives
	x	0.067282918	0.067282386		x	0.023478708	0.023479086
C_1	y	-0.105823768	-0.105823357	H_3	y	0.063189104	0.063189051
	z	-0.038120401	-0.038122458		z	-0.004080645	-0.004080662
	x	-0.023618805	-0.023618761		x	-0.067142821	-0.067142787
O_2	y	0.037092253	0.037092194	H_4	y	0.005542411	0.005542200
	z	0.046339681	0.046341970		z	-0.004138636	-0.004138668
					Job time:	9.7	143.6

Table 4.1: Value (Hartrees/Bohr) of the analytical and numerical geometrical gradient for HF, H_2O and $H_2C = O$. Standard deviation from numerical data are respectively σ_{HF}=5.8 $\cdot 10^{-8}$, σ_{H_2O}=8.4$\cdot 10^{-7}$, $\sigma_{H_2C=O}$=9.18$\cdot 10^{-7}$. Computational time required is also reported in seconds.

4.1 which clearly shows the effects. We can see that for numerical gradients calculations the computational time growths exponentially with respect to the number of nuclei of molecules, whereas a linear trend with very flat slope is obtained with the analytical formulation. In the case of the gradients calculation for butadiene the analytical algorithm is 33 times faster then the numerical one.

4.1.2 Equilibrium geometry of *s-trans* 1,3-butadiene

We first consider the effect of the pressure on the equilibrium geometries of the s-trans 1,3-butadiene. Table 4.4 reports the values of the pressure, and of the solutesolvent Pauli repulsion energy, experienced by *s-trans* 1,3-butadiene as a function of the volume of the cavity and of the corresponding scaling factor *f*. The numerical results show that as the scaling factor f (and parallel the cavity volume) is reduced, the Pauli repulsion and the pressure rapidly increase. A regime of very high pressure condition ($p > 10$ GPa) is reached for scaling factor values of 0.8 and 0.7, corresponding to a pressure of 13.5 and 22.0 GPa, respectively.

The equilibrium bond lengths of *s-trans* 1,3-butadiene as a function of the pressure are

C_2H_4		Analytical	Numerical			Analytical	Numerical
Atom	Coord.	derivatives	derivatives	Atom	Coord.	derivatives	derivatives
	x	0.019195958	0.022501133		x	-0.106489280	-0.109792296
C_1	y	0.016094584	0.015617235	H_4	y	-0.018581302	-0.018101122
	z	0.004716112	0.008131694		z	-0.055382461	-0.058799761
	x	-0.020197736	-0.020196926		x	-0.086303296	-0.086303898
C_2	y	0.014807843	0.014808221	H_5	y	-0.000547738	-0.000547740
	z	-0.007914429	-0.006256900		z	0.021209752	0.021209725
	x	0.090404351	0.090404926		x	0.103390003	0.103389885
H_3	y	0.002486642	0.002486629	H_6	y	-0.014260028	-0.014260259
	z	-0.018369722	-0.018369704		z	0.055740747	0.054084247
					Job time:	11.6	241.6
C_3H_6		Analytical	Numerical			Analytical	Numerical
Atom	Coord.	derivatives	derivatives	Atom	Coord.	derivatives	derivatives
	x	-0.000465582	-0.000465576		x	0.037080375	0.037116892
H_1	y	-0.003237798	-0.003237785	C_6	y	0.000298386	0.000289743
	z	0.092953458	0.092953215		z	0.094715343	0.094715961
	x	-0.082370043	-0.082370622		x	-0.009594428	-0.009594087
C_2	y	0.143354636	0.143354359	H_7	y	0.000463094	0.000463116
	z	-0.055622990	-0.055622718		z	0.010010968	0.010010970
	x	0.081545682	0.081546232		x	0.000976770	0.000939129
H_3	y	-0.042860976	-0.042860952	H_8	y	0.011813019	0.011820814
	z	0.024541191	0.024541147		z	0.001030234	0.001030440
	x	-0.048366927	0.048366485		x	0.004117856	0.004118613
C_4	y	-0.118033528	-0.118033319	H_9	y	-0.008039223	-0.008038500
	z	-0.136813725	-0.136814415		z	0.003173129	0.003173115
	x	0.017076297	0.017075852				
H_5	y	0.016242390	0.016242411				
	z	-0.033987607	-0.033987633				
					Job time:	20.6	485.9

Table 4.2: Value (Hartrees/Bohr) of the analytical and numerical geometrical gradient for C_2H_4 and C_3H_6. Standard deviation from numerical data are respectively $\sigma_{C_2H_4}$=1.7 $\cdot 10^{-3}$, $\sigma_{C_3H_6}$=1.0$\cdot 10^{-5}$. Computational time required is also reported in seconds.

C_4H_6 Atom	Coord.	Analytical derivatives	Numerical derivatives	Atom	Coord.	Analytical derivatives	Numerical derivatives
	x	-0.015638198	-0.015638190		x	-0.007165284	0.004344540
H_1	y	0.000001070	0.000001083	C_6	y	-0.000385113	-0.000385137
	z	-0.000849680	-0.000850242		z	0.015034866	0.031480184
	x	0.053629589	0.053629547		x	-0.005186696	-0.016703862
C_2	y	-0.000045653	-0.000045604	H_7	y	0.000456322	0.000456271
	z	-0.053356436	-0.058263881		z	0.022662191	0.011142406
	x	0.000811813	0.000812209		x	0.023912354	0.024913218
H_3	y	-0.000002391	-0.000002421	C_8	y	-0.000288529	-0.000288499
	z	-0.001424562	-0.001424634		z	0.010719145	0.010719356
	x	0.006754824	0.006754802		x	-0.032879904	-0.032879948
C_4	y	-0.000042904	-0.000042894	H_9	y	0.000052413	0.000052352
	z	-0.077517674	-0.077518144		z	0.018771432	0.018771746
	x	-0.016970111	-0.016970511		x	-0.007268385	-0.007268935
H_5	y	0.000029683	0.000029679	H_{10}	y	0.000225103	0.000225172
	z	0.031337167	0.031337178		z	0.034623551	0.034623635
					Job time:	45.2	1494.2

Table 4.3: Value (Hartrees/Bohr) of the analytical and numerical geometrical gradient for C_4H_6. Standard deviation from numerical data are respectively $\sigma_{C_4H_6}$=4.8 $\cdot 10^{-3}$. Computational time required is also reported in seconds.

f	V(Å^3)	p (GPa)	G$_{rep}$(Kcal/mol)
-	-	0.000	0.0
1.40	137.46	0.10	0.58
1.20	99.11	0.81	3.89
0.91	55.07	6.08	22.40
0.80	42.16	13.39	44.10
0.72	33.49	21.58	67.20

Table 4.4: Values of the cavity volume, V (Å^3), of the pressure, p (GPa) and of the repulsion energy G$_{rep}$ (kcal/mol) for trans 1,3-butadiene, as a function of the shrinking cavity factor f. Results refer to the PCM/DFT/B3LYP/6-31G(d,p) level. See text for the details of the PCM cavity.

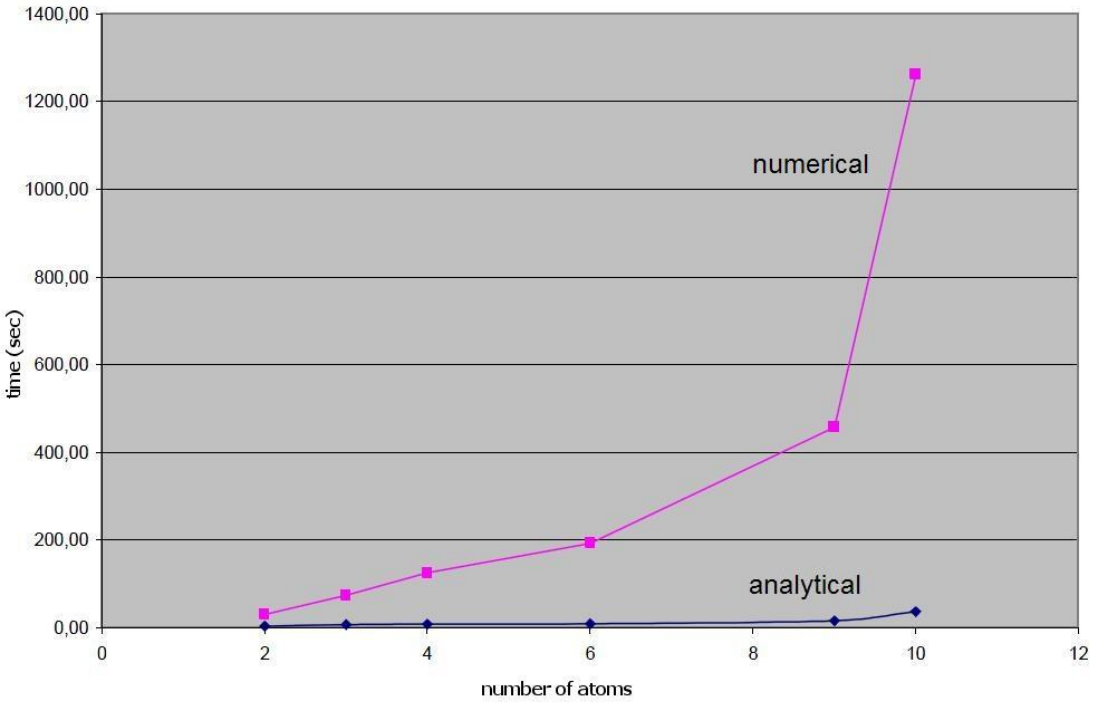


Figure 4.1: Analytical vs Numerical Gradient - Comparison between numerical (square) and analytical (rhomboid) computational time of the molecular gradient related to the upper introduced chemical systems

shown in Table 4.5 The data show a general shortening of the bond lengths with the increase of pressure.

p (GPa)	r (C=C)	r (C-C)	r (C-H_a)	r (C-H_b)	r (C-$H_{b'}$)
0.00	1.3409	1.4576	1.0912	1.0886	1.0866
0.10	1.3412	1.4578	1.0914	1.0887	1.0858
0.81	1.3411	1.4579	1.0888	1.0874	1.0854
6.08	1.3394	1.4541	1.0855	1.0800	1.0823
13.39	1.3371	1.4517	1.0826	1.0815	1.0779
21.58	1.3342	1.4489	1.0787	1.0785	1.0758

Table 4.5: Bond lengths (Å) of trans 1,3-butadiene as a function of the pressure, p (GPa). Results refer to the PCM/DFT/B3LYP/6-31G(d,p) level. CH_a denotes the inner CH bonds, while CH_b and $CH_{b'}$ denote the outer CH bonds, respectively, in cis and trans position with respect to the inner CH bonds. Bond angles do not show significative variations with respect to the value of the isolated molecules, and have not been reported.

The mean value of the C-H bond lengths decreases to 0.013 Å at 21.2 GPa with respect to its vacuum value; this corresponds to a bond compressibility of 6.1x10^{-4} Å GPa^{-1}. The mean value of the single (C-C) and double (C=C) bonds decrease, respectively, of 0.0087 Å and 0.0067 Å (at 21.2 GPa); these correspond to bond compressibility values of 4.1x10^{-4} Å GPa^{-1} and 3.2x10^{-4} Å GPa^{-1}, respectively. The ratio of the C-C/ C=C compressibility is then 1.28. These compressibility can be compared with those estimated from the Carr.Parrinello simulation by Cardini and co-workers (54) on 1,3- butadiene at very high pressure. In such a study, the mean intra molecular carbon−carbon distances can be obtained from the maxima of the carbon−carbon pair radial distribution functions $g(r)$, and the corresponding bond compressibility can be estimated from the shift of the maxima as a function of the pressure. The simulation of 1,3-butadiene at the density of 1.87 g cm^3 gives a value of pressure of about 40 GPa, to which corresponds a variation of 3.2x10^{-2} Å and 2.9x10^{-4} Å , respectively, for the double and single CC bonds. This gives an estimation of 8.0x10^{-4} Å GPa^{-1} and 7.2x10^{-4} Å GPa^{-1}, respectively, for C-C and C=C bonds, compressibility, corresponding to a ratio of CC/C=C compressibility 1.1, in reasonable agreement with our results.

The shortening of the bond lengths of 1,3-butadiene may be interpreted, within the Born-Oppenheimer approximation, as a consequence of the deformation of its electronic charge density, due to the Pauli repulsive interaction with the environment. In fact, this interaction pushes the tails of the peripheral electronic distribution toward the internal regions of the molecule.

The corresponding depletion of electron density in the peripheral region determines, according to the electrostatic (HellmannFeynman) theorem, a force on the nuclei in the direction of the internal region of the molecule, with a general shortening of the bond lengths (239).

These effects on the electronic density of *s-trans* 1,3-butadiene are clearly shown in Fig. 4.2, where the isosurfaces of the electron density variations as a function of two representative pressure values are reported. Another interpretation of the shortening of the length of CC bonds upon confinement can be given in terms of variation of the orbital energies in response to the molecular confinement (240).

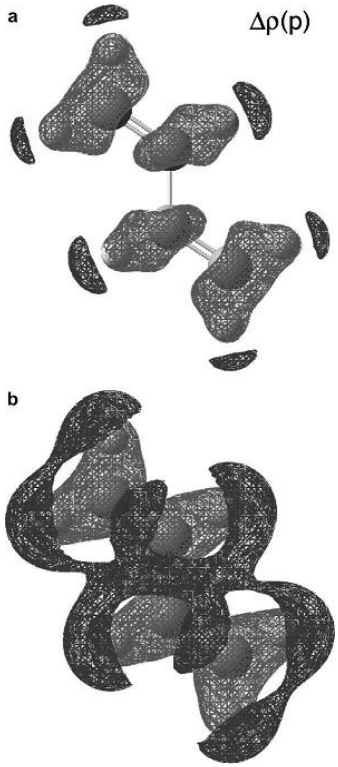


Figure 4.2: Isosurface variation - Shifts in the electron density of trans 1,3-butadiene due to the pressure, p, $\Delta\rho(p)=\rho(p)-\rho(0)$. Density difference isosurfaces are at $\Delta\rho(p)=0.00035$ e/a_0^3; deep gray surfaces enclose volumes in which electron density decreases: (a) isosurfaces at p = 4.4 GPa, (b) isosurfaces at p=27.2 GPa

4.1.3 Conformational energies of 1,3-butadiene

We now go ahead to consider the effect of the pressure on the relative conformational energies of the *s-trans*, *s-cis* and skew conformers of 1,3-butadiene. As the *s-trans*, the s-cis conformer

also is planar, while the skew is cisoid but not planar. Various types of experimental gas phase studies have shown that *s-trans* conformation is the most stable. A small amount of the skew conformers is present in equilibrium with the major conformer (241).

Theoretical calculations have shown that the *s-trans* conformer is 2 5 kcal/mol lower than either the s-cis or skew conformers. Most high-level calculations favor the skew conformation over the planar cis, but the energy differences found are quite small (241).

In Table 4.6, we report the relative conformational energies $\Delta G_{el-rep} = G_{el-rep}(X) - G_{el-rep}(s-trans)$, X = s-trans, cis, skew as a function of the pressure. These data refer to the conformers in their equilibrium geometries, and neglect both the zero-point and the cavity formation energy contributions. For the whole range of pressure (027.2 GPa), *s-trans* is the most stable conformation, although the pressure slightly stabilizes the other two conformers: at 27.2 GPa the skew and s-cis conformers are stabilized, respectively, by 0.6 kcal/mol and 0.5 kcal/mol, with respect to the vacuum. However, the effect of the pressure on the conformational energies is not monotonic and both conformers show a local minimum at 7.8 GPa. This ancillary aspect is worth to be considered in more detail in a future work. The overall stability of the *s-trans* conformer is in agreement with a recent Raman study of 1,3-butadiene at high pressure (242). Although high pressure has a little influence on the conformations of 1,3-butadiene, due to the little variation of the molecular volume, in general, its effect on more conformation flexible molecular systems may be of relevance. The reader interested to this aspect is referred to the papers by Whalley (243) and by Walker and co-workers (244) and references therein.

p (GPa)	ΔG_{el-rep}(X)(Kcal/mol)		
	s-trans	skew	s-cis
0.0	0.0	3.53	3.87
4.4	0.0	3.89	3.93
7.8	0.0	3.43	3.39
14.3	0.0	3.48	3.76
27.2	0.0	2.91	3.39

Table 4.6: PCM/DFT/B3LYP/6-31G(d,p) relative energies of the conformers of 1,3- butadiene ($\Delta G_{el-rep} = G_{el-rep}(X) - G_{el-rep}(s-trans)$, X = s-trans, cis, skew) as a function of the pressure.See text for the details of the calculations.

4.2 Computational results in transition metal applications

4.2.1 Exited state properties of two luminescent silver complexes

In Table 4.7 are reported first calculated transitions, respectively for $[Ag(PPh_3)_2(phpty)]^+$ and $[Ag(PPh_3)_2(phmepty)]^+$, with respective absorption wave length, oscillator strengths, mean MOs composition and hypothetical assignment.

The rendering of MOs at 0.02 iso-dansity (Fig.4.3) is a good tool for understand principal meaning of transition involving excited states reported in Table 4.7. HOMO of both complex is metal,phosphorus- ligand localized with small thio-ligand contribution. LUMO is, in each system, poorly metal localized and strongly diffuse on pyridine side of thio-ligand. This is true also for LUMO+1 in $[Ag(PPh_3)_2(phmepty)]^+$ but not in $[Ag(PPh_3)_2(phpty)]^+$ where high probability to find electrons on the metal was calculated.

If we study thio-amidic LUMO and LUMO+1 molecular orbitals, we find that only in the first complex there is a no zero electron density on N,S atoms, whereas for the second wave function is confined on the pyridine ring.

Other MOs involved in excitation process such as HOMO-1, HOMO-2 and LUMO+2 are very interesting for a good comprehension of these mechanism but for the preliminary character of these results we will say no more about them.

For the first complex we can say that most representative transition regards extra ligand (XLL) charge transfer from PPh_3 group to *phpytu* and inter-ligand (ILL) into conjugate system phenyl-pyridine. The effect of metal to ligand charge transfer (MLCT) at 272.31 nm is substantially covered by a more intense contribution of the transition from HOMO-1 $\rightarrow$ LUMO+2 and HOMO-2 $\rightarrow$ LUMO.

For the second complex we can obtain similar conclusion by MOs analysis, but with more consideration can be done for LUMO and LUMO+1 that differs substantially from analogue in the first complex. In fact MOs reported in Fig.4.4, are characterized by no metal localization for *b)* on the contrary that for *a)*. This comport a more significant contribution by MLCT for all transition within XLL and ILL still present.

Regard zero electron density on N and S atoms in the second complex against what appends for the first one, we can think that methylene ($-CH_2$) group, situated in pyridine side of the ligand, acts against low energy resonance onto S-Ag bond with consequent less de localization.

In conclusion we can attribute principal mechanism in UV/vis absorption to charge transfer from phosphine to our two ligand and transfer between them. It is no easy to distinguish

State	λ (nm)	Osc.	Composition	Assignment
1 ^{1}A	310.90	0.0651	HOMO-1 → LUMO(0.44)	MLCT
			HOMO → LUMO(0.46)	XLL
2 ^{1}A	302.82	0.0211	HOMO-1 → LUMO(0.43)	MLCT
			HOMO → LUMO(0.49)	XLL
3 ^{1}A	294.54	0.0105	HOMO → LUMO+1(0.35)	XLL
			HOMO → LUMO+2(0.58)	
4 ^{1}A	283.40	0.0822	HOMO-1 → LUMO+1(0.35)	XLL+ILL
			HOMO-1 → LUMO+2(0.53)	
5 ^{1}A	281.73	0.0021	HOMO → LUMO+1(0.57)	XLL+ILL
			HOMO → LUMO+2(0.34)	
6 ^{1}A	272.31	0.0584	HOMO → LUMO(0.27)	MLCT+XLL
			HOMO-1 → LUMO+2(0.33)	ILL
			HOMO-2 → LUMO(0.46)	XLL+ILL

State	λ (nm)	Osc.	Composition	Assignment
1 ^{1}A	351.43	0.0028	HOMO → LUMO(0.70)	MLCT+XLL
2 ^{1}A	337.58	0.0075	HOMO-1 → LUMO(0.70)	MLCT+XLL
3 ^{1}A	313.48	0.0033	HOMO → LUMO+1(0.67)	MLCT+XLL
4 ^{1}A	305.14	0.0051	HOMO-1 → LUMO+1(0.52)	MLCT+XLL
5 ^{1}A	301.81	0.0199	HOMO-2 → LUMO(0.46)	MLCT+ILL
				XLL
6 ^{1}A	294.50	0.0020	HOMO-1 → LUMO+1(0.36)	MLCT+XLL
			HOMO-1 → LUMO+2(0.31)	XLL
			HOMO → LUMO+2(0.44)	XLL+ILL

Table 4.7: Computational results of excitation energy (λ,nm), mean MOs composition of relatives excited states and hypothetical assignment in absorption transition of $[Ag(PPh_3)_2(phpty)]^+$ and $[Ag(PPh_3)_2(phmepty)]^+$.

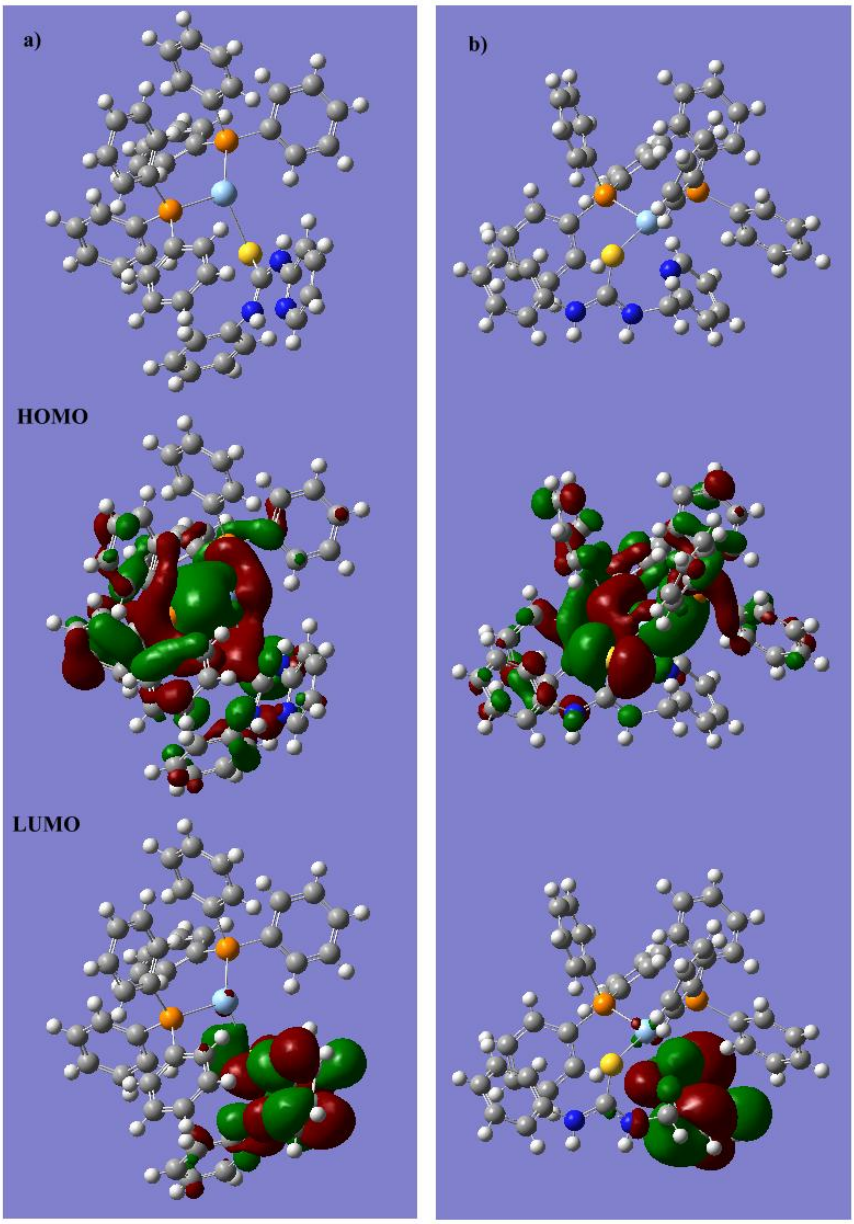


Figure 4.3: Molecular Orbitals Comparison - Frontier MOs diagram calculated respect to lowest excitations obtained with TDDFT theory for a) $[Ag(PPh_3)_2(phpty)]^+$ and b) $[Ag(PPh_3)_2(phmepty)]^+$

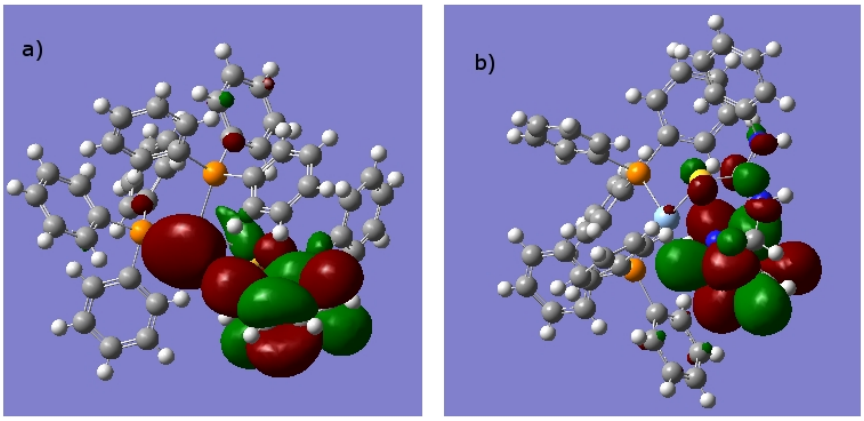


Figure 4.4: Molecular Orbitals Comparison 2 - LUMO+1 MOs diagram calculated respect to lowest excitations obtained with TDDFT theory for a) $[Ag(PPh_3)_2(phpty)]^+$ and b) $[Ag(PPh_3)_2(phmepty)]^+$

metal to ligand charge transfer from others because more detailed theoretical and experimental analysis should be done. Reproduction of UV/vis spectra based on energy of excited states, oscillator strength and a gaussian convolution sum, that we had analyzed for this elaboration shown that most of discussed transition broad into two band principally represented by strength oscillator due to XLL and ILL.

4.2.2 Computational analysis into hydrogen transfer Ru assisted

The optimized structures of the pre catalyst, reaction intermediates, and transition states are collected in Fig. 4.5 and 4.6.

The ring-closure process leading to **1a** (Fig. 4.7) was studied by DFT calculations in vacuo and in iPrOH solution on the basis of Equation 4.1. Here the organometallic cation $Ru[(\kappa^2P,N)PNH_2](p\text{-cymene})Cl^+$ is indicated as III. All the reagents present in the reactant solution at the onset of the reaction were considered to be present in the initial step. The starting reactants were located at 0.0 kcal/ mol energy. This means that neutral complex **1** is located at 0.0 kcal/mol in the energy profile diagram depicted in Fig. 4.8.

$$1 \rightarrow III + Cl^- \tag{4.1}$$

Following we are going to describe the energy profile obtained with the calculations performed at the BS-IV level (see section 3.2.2.2). The complete exploration of the energy surface

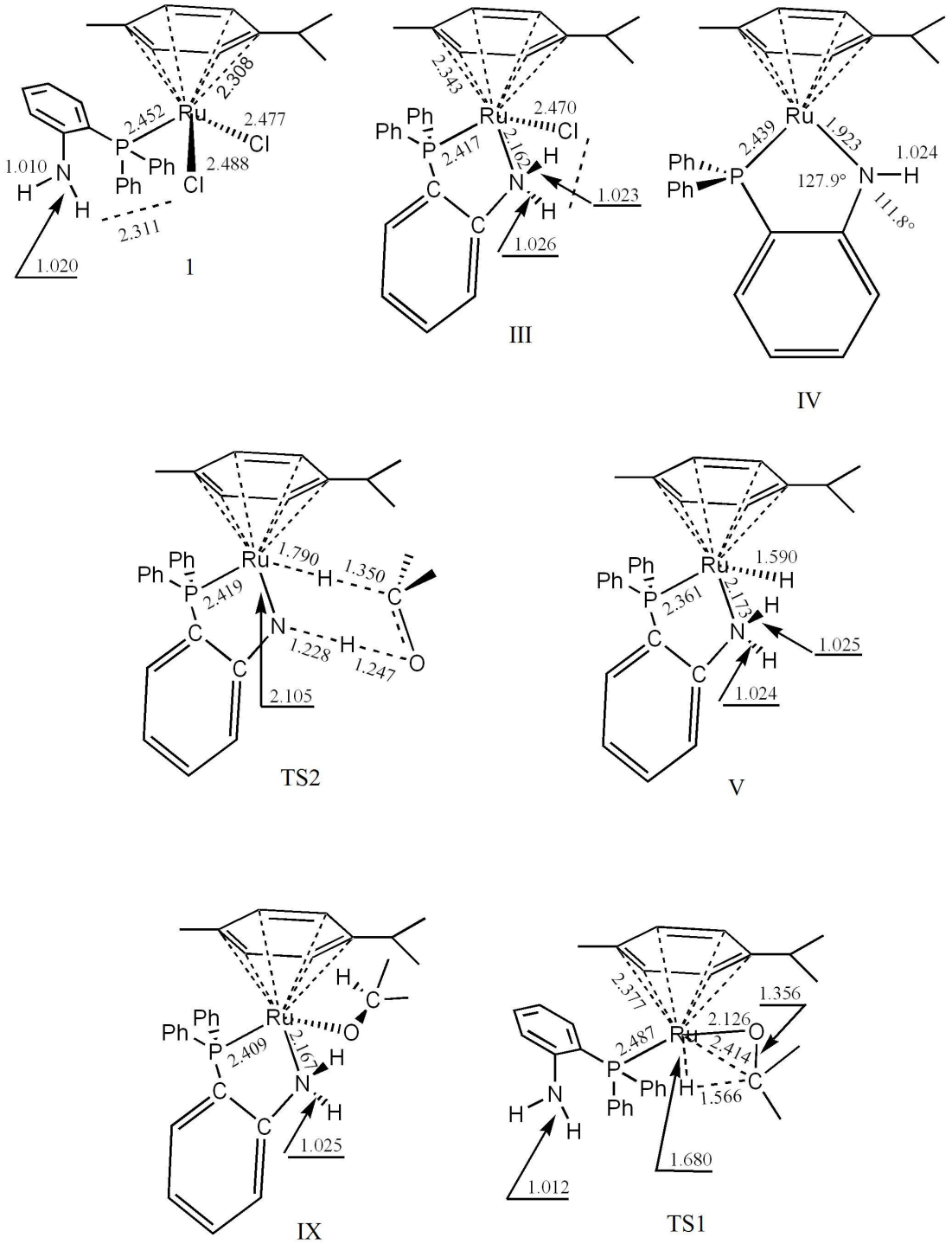


Figure 4.5: Geometry structure - Optimized geometries at B3LYP/BS-0 level of stable intermediate complexes and transition states, related to Paths I and II. For simplicity, hydrogen atoms are omitted except for hydride and aminic functionalities, whereas the phenyl rings of the phosphane moieties are drawn as single atoms

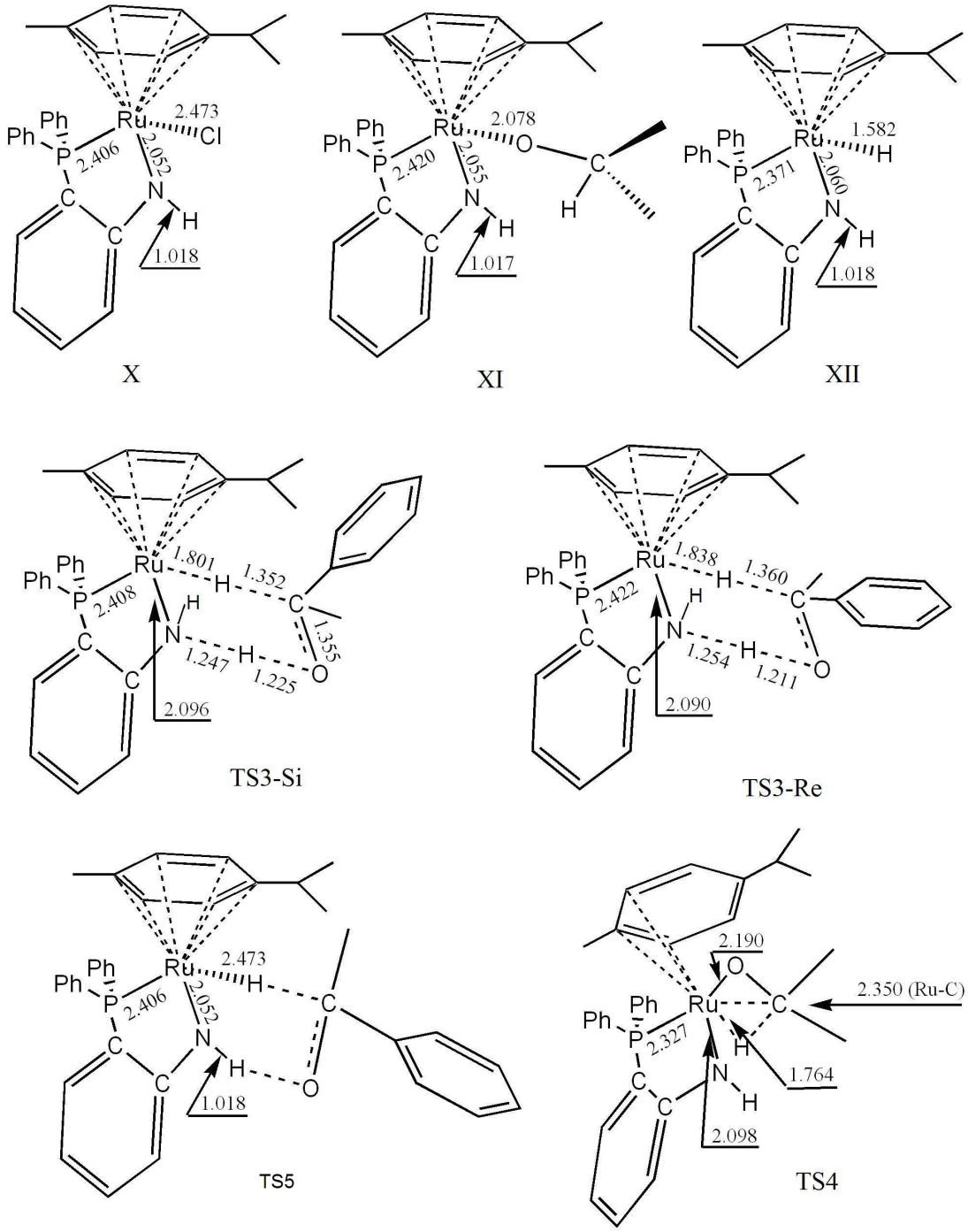


Figure 4.6: Geometry structure continue - Optimized geometries at B3LYP/BS-0 level of stable intermediate complexes and transition states, related to Paths I and II.For simplicity, hydrogen atoms are omitted except for hydride and aminic functionalities, whereas the phenyl rings of the phosphane moieties are drawn as single atoms

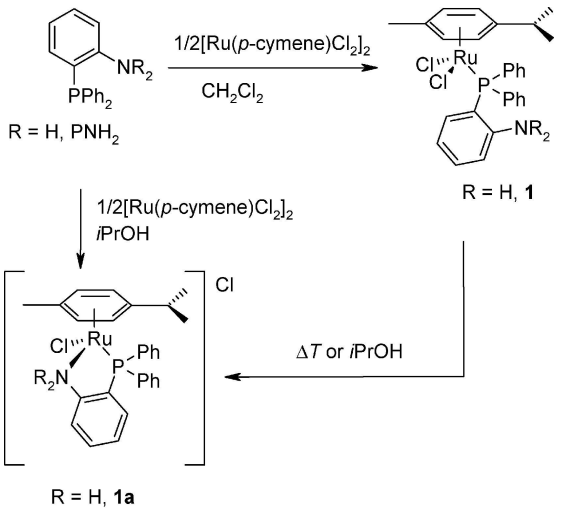


Figure 4.7: Ring Closure - Isomerization reaction scheme of **1** in **1a**

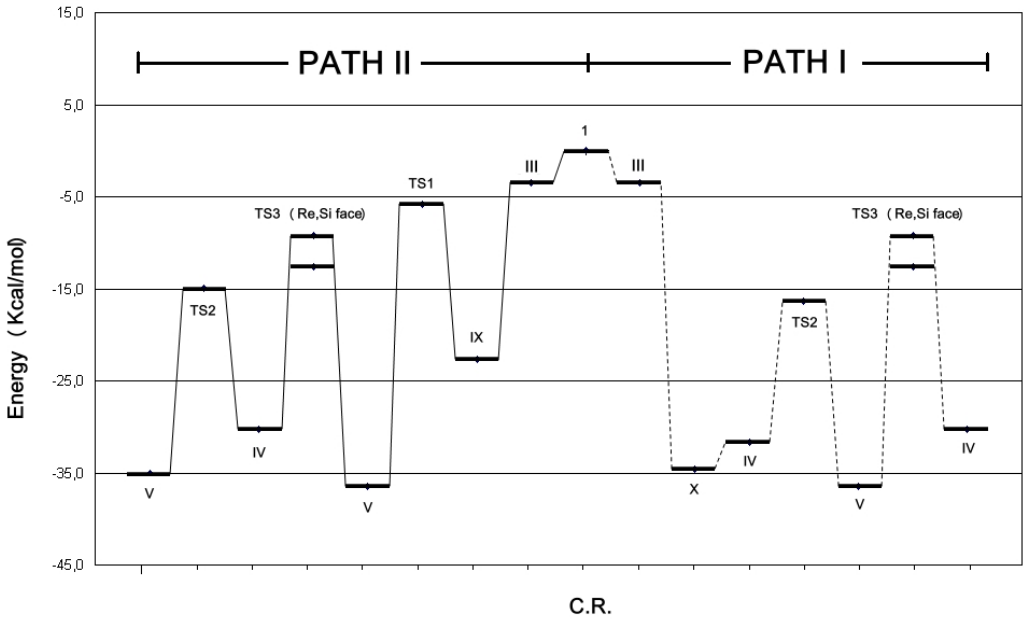


Figure 4.8: Energy Profile - Comparison of the energy profiles of Paths I (on the right) and II (on the left). All the numbered reaction steps are reported as ΔG with respect to **1** and other reactants (see Appendix Reaction List 8). Only the most important catalytic species are labeled in the graph

in function of the different basis sets is reported in Appendix Table 7.11. The calculations performed *in vacuo* show an endothermic profile (87.3 kcal/mol) that becomes exothermic once iPrOH is introduced into the model (3.5 kcal/mol, Fig. 4.8. This is in agreement with the experimental observation that the chelation of the PNH_2 ligand occurs easily in iPrOH, where cation III and anion Cl^- are certainly stabilized, whereas in less-polar solvents the breaking of the RuCl bond is unlikely, and it occurs appreciably only at temperatures ≥ 50 C.

As can be seen in Fig. 4.5, the chelation leads to a shortening of the RuP bond, from 2.452 Å in **1** (2.376 Å in the X-ray structure) to 2.417 Å in III, in accordance with the solid-state structure data of chelate complexes **1a** where the Ru-P bond lengths are 2.3072, 2.296, and 2.310 Å respectively. Moreover, an intra molecular hydrogen bond between the axial aminic proton and the coordinated chloride ligand is also found (2.331 Å).

Now cation III can react with $iPrO^-$ by Path I (Fig. 4.9) or by Path II (Fig. 4.10). The energy profile of Path I (Fig. 4.8) indicates that the amine deprotonation (Fig. 4.9, step (i)) occurs exothermally with a release of energy of 28.1 kcal/ mol. Step (i) can be divided into two subsequent steps: amine deprotonation with formation of the transient neutral complex Ru[(κ^2 P,N)PNH](p-cymene)Cl(X; Fig. 4.9, step (ii)) and chloride dissociation [Fig. 4.9, step (iii)]. The energy difference that separates X and IV is low (2.9 kcal/ mol, Fig. 4.8), pointing out a facile dissociation of the chloride ligand (Any attempts made to isolate the neutral complex Ru[(κ^2 P,N) PNH](p-cymene)Cl (X) were unsuccessful (235)). Structural analysis of the optimized geometry of IV shows an evident shortening of the Ru−N bond (1.923 Å) with respect to the values found in the complexes crystallographic characterized, in agreement with a substantial amide character of the coordinated nitrogen.

This is confirmed by the comparison of the calculated RuNH bond angles in IV and V, which are 120.2 ^{0} and 109.9^0, respectively. On the contrary, the calculated Ru-P bond length is sensibly longer in IV with respect to those found in the X-ray structures of **1a** and to that calculated for V, passing from 2.408 to 2.305 Å(averaged value) to 2.361 Å, respectively. This is a new and never reported structural distortion for the classical bifunctional catalyst. As Yamakawa et al. has been reported (134; 142; 143) one of the most important structural distortion of the bifunctional ring bonded to the metal is the drastic bond length shortening of both the HN-M and O-M. In particular, this remark, has been used to justify the unexpected stability of the 16e- configuration which is supposed to be very difficult to isolate, if no electron density compensation occurs. This elongation could especially explain the difficulty to isolate [IV]Y (Y = Cl^- or PF_6^-) as a solid, as mentioned above (235). In Path II, the first event is the formation

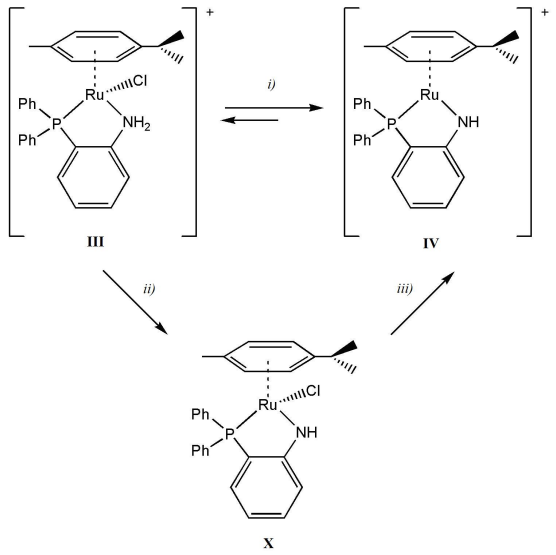


Figure 4.9: Amine Deprotonation - deprotonation of the amine functionality with formation of IV. Conditions: (i) $iPrO^-$, iPrOH, Cl^-; (ii) $iPrO^-$; iPrOH; (iii) Cl^-

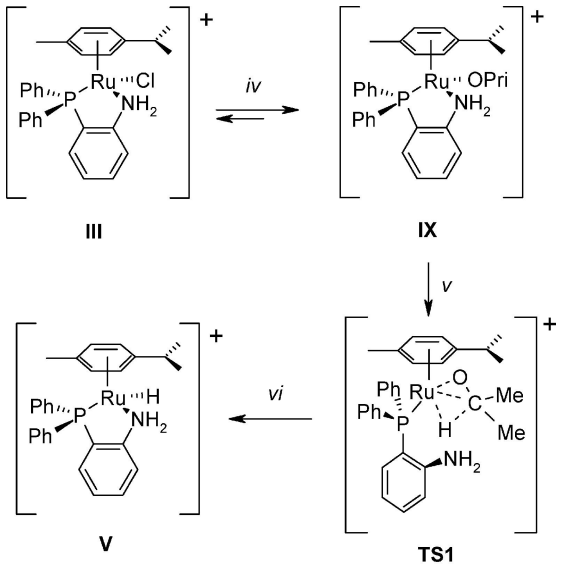


Figure 4.10: Path II - Reaction steps for Path II. Conditions: (iv) $iPrO^-$, Cl^-; (v) TS for the -hydrogen elimination step; (vi) Me_2CO

of isopropoxide intermediate IX, as depicted in Fig. 4.10 (step iv). This step is characterized by a pronounced exothermicity (19.2 kcal/mol, Fig. 4.8).

The subsequent β-hydrogen elimination occurs through TS1, where the amine function is not coordinated to ruthenium allowing the formation of an agostic interaction between the α-hydrogen of the alkoxide and ruthenium (Fig. 4.10, step (v)). This step is quite costly in energy requiring 16.9 kcal/mol with respect to IX. The completion of the β-hydrogen elimination leads to hydride V and acetone (Fig. 4.10, step (vi)), with a gain in energy of 30.7 kcal/mol with respect to TS1. The high exothermicity of the last step implies that once the hydride species is formed the reversed process is strongly unlikely.

From an energetic point of view Path II cannot be considered unproductive, as the required energies do not appear inaccessible. However, the absence of experimental evidences about alkoxide species, as well as the fact that the observed catalytic activities do not increase with base, suggest that Path I is definitely more likely.

Once hydride V is formed, the process evolves through Fig. 4.11. Thus, 16e species IV interacts with iPrOH through TS2 (Fig. 4.11, step (vii)). This step requires 15.3 kcal/mol and leads to V and acetone with an exothermic gain of 20.2 kcal/mol. The hydroxy proton is transferred to the nitrogen atom, whereas the C−H proton is transferred to ruthenium (Fig. 4.11, step (viii)), and the whole oxidation of iPrOH to acetone gives -4.9 kcal/mol.

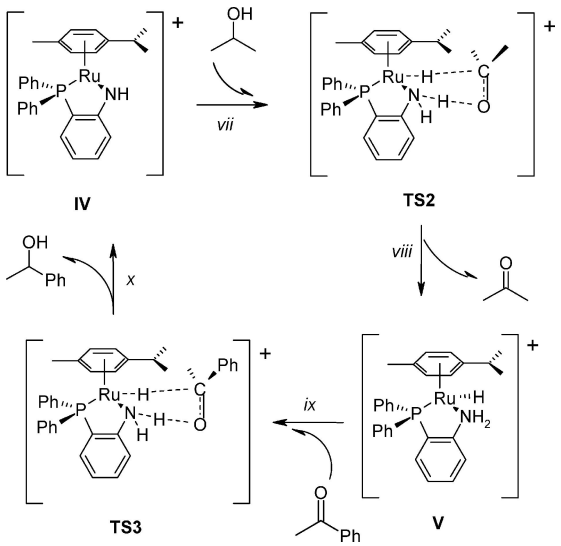


Figure 4.11: Hydrogen transfer cycle - OSM in action for the reduction of acetophenone catalyzed by III. Intermediate X has been omitted

Now V interacts with acetophenone through the pericyclic transition state TS3 (Fig. 4.11, step (ix)), where the oxygen atom forms a hydrogen bond with the axial amine proton of V, whereas the hydride functionality interacts with the carbonyl carbon atom of the ketone.

For the construction of TS3 we fixed the *S* configuration of ruthenium and we approached the ketone by two different enantiofaces, *Re* and *Si*. DFT calculations evidence two different energetic profiles, and the *Si* was favored over the *Re* face (23.8 and 27.2 kcal/mol with respect to the previous step). This energy differences (3.4 kcal/mol) arise from specific CH(p-cymene)π(phenyl) interactions that are possible only when the ketone approaches ruthenium from the Si face. This enantioface discrimination was invoked at the origin of the observed high enantioselectivity found with Noyoris catalysts. In particular we observed that the *Si* face approach lead to an evident shortening of the distance between the phenylic ring of the keton and the p-cymine fragment, easily displayed in Fig. 4.12 , where the distance between hydrogen (on the p-cymine) and carbon (on the phenylic group) became 2.89 Å, lower then the sum of the respective Van der Waals radii (2.9 Å) which is obviously not possible on the *Re* face approach (Fig. 4.13). A complete molecular orbital (MO), frequency (FQ), Atom

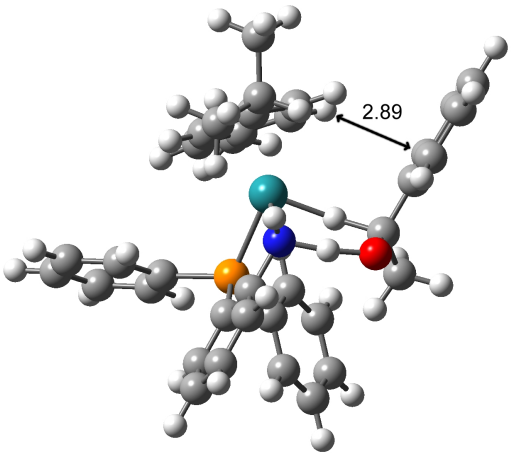


Figure 4.12: pro-Si approach - The interaction between keton and the catalyst on the pro-Si face is characterized by a forced drawing near of the phenyl group toward the p-cymene

in Molecule (AIM), as well as a Natural Bond Orbital (NBO) analysis has been carried out to characterize this transition state and the C-H specific interaction. The TS3-*Re* and TS3-*Si* was proven to have respectively i960 and i940 cm^{-1} single imaginary frequency (These and more others frequencies calculation are collected in Appendix Table 7.12). In Fig. 4.14b is

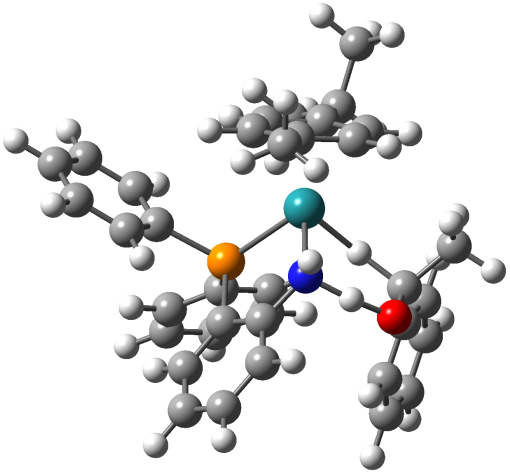


Figure 4.13: pro-Re approach - In the case of a pro-Re approach there is no way to force the interaction between C atoms of the phenyl group and H of p-cymene. In this case no specific interaction like C-Hπ can stabilize the related transition state

pictured the normal mode vector relative to TS3-Si where the hydride hydrogen and the proton synchronously move on the keton and it is also possible to see the higher amplitude of N-H displacement respect to that the Ru-H does.

It is also possible to note the evident hybridization change from $sp^2 \rightarrow sp^3$ of the carbonyl group and in contraposition the flattening of the aminic group ($sp^3 \rightarrow sp^2$).

The MOs analysis (Fig. 4.14a) show that the *s* orbital of the hydride is oriented toward the perpendicular *p* shaped orbital of the carbonyl carbon but without any direct overlap caused of a nodal plane between the two atoms. On the other side of the ring, the aminic hydrogen experiment the same kind of nodal plane, which describe a typical HOMO shape in the hydrogen bond interaction.

To clarify the CH interaction we performed a NBO analysis on the TS3-Si, on the isolated ketone as well as for the isolated p-cymene. It is clear (Fig. 4.14c) that the aromatic hydrogen (charge = +0.2220 u.a.) of the p-cymene group experiment an increasing of their elettrophylity once the coordination to the metal occurs (+0.2726 u.a.) in complete agreement with what is reported in literature (245). On the other hand a completely opposite trend interests the aromatic carbon of the phenylic group of the keton. In particular the approach of negative hydride on the carbonyl group (this decrease the electrophylity of C7 from +0.7074 to 0.4384) implies a strong increasing in the nucleophylity of the ortho/para carbons moving C3 charge from -

0.1650 in the free acetophenone to -0.2275 in to the transition state TS3-Si. This interesting re distribution of charge due to the concerted transition state is not a prerogative of TS3-Si in fact very similar behavior occurs in the TS3-Re but, in this case, the distance between the p-cym ring and the phenylic group doesn't allow a specific CH interaction.

We confirm, in our system, that the enantioface discrimination comes from this specific interaction which has been invoked to justify the observed high enantioselectivity found with Noyori type catalysts (142; 143; 144; 145). However, because V is expected to be present

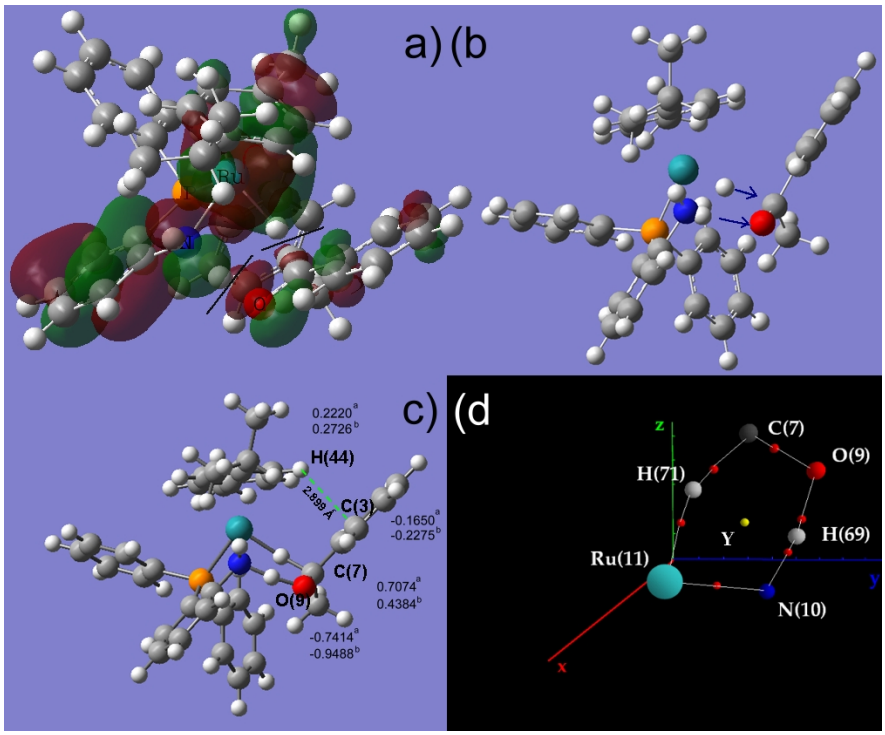


Figure 4.14: Analysis on the transition state TS3 - a) HOMO rendering of the transition state where the two characteristic nodal plane are observable; b) picture of the normal mode vector of the transition state related to the hydrogen transfer; c) Mulliken^a and NBO^b population of the most important atoms involved into the hydrogen transfer; d) Bond critical point obtained by AIM analysis on the pericyclic transition state

as a racemate in solution, the choice of a precise configuration of the metal and of a precise enantioface of the ketone is totally arbitrary (no asymmetric induction was observed with **1a**). Finally in Fig. 4.14c and Table 4.8 are shown the results related to the AIM analysis. The picture is simplified and just represents the pericyclic ring which characterize the transition state and in particular the related bond critical point (BCP). The BCP founded in the middle

of the ring is the characteristic (3,+1) critical point which is always detected in ring structures. In table we report the electron density ρ and its Laplacian $\nabla^2\rho$ which is in relation with the chemical nature of the bond where the critical point has been found.

BCP	ρ	$\nabla^2\rho$
N(10)-H(69)(3,-1)	0.1694	0.0661
H(69)-O(9)(3,-1)	0.1577	0.0368
O(9)-C(7)(3,-1)	0.2857	0.1363
C(7)-H(71)(3,-1)	0.1343	0.0236
H(71)-Ru(11)(3,-1)	0.0742	-0.0393
Ru(11)-N(10)(3,-1)	0.0980	-0.0927
Y(3,+1)	0.0166	-00219

Table 4.8: A.I.M. analysis of TS3 structure. We report bond critical point (3,-1) located on bond path with relative electronic density ρ and Laplacian $\nabla^2\rho$. Ring critical point(3,+1) was also characterized.

The simultaneous transfer of the amine proton and hydride ligand to the ketone leads to the formation of IV and 1-phenylethanol (Fig. 4.11, step (x)).The possibility that TS3 could involve an iPrOH molecule, as described in J. W. Handgraaf, E. J. Meijer, J. Am. Chem. Soc. 2007, 129, 3099, was investigated by *ab initio* calculations. In particular, we searched saddle points after the introduction of one explicit molecule of iPrOH bridging acetophenone and the NH_2 functionality of V through hydrogen bonds. This led to a stable transition state with only one imaginary frequency describing synchronous hydrogen transfer from NH_2 and RuH functionalities. Preliminary calculations show a sensible destabilization in energy of TS3. The 16e species IV is 17.6 kcal/mol more stable than TS3 (considering the Si face of acetophenone) and the reduction of acetophenone to 1-phenylethanol needs 6.2 kcal/mol. Thus, the ΔG value of the entire process of oxidation of iPrOH and reduction of acetophenone is 1.3 kcal/mol.

As regards Path III, the optimized geometries of the reactant intermediates and transition states are collected in the Fig. 4.6. The ground-state intermediates are depicted in Fig. 4.15. Here we have considered that the initial deprotonation of **1a** gives X, iPrOH, and Cl^-. Compound X reacts with a second $iPrO^-$ ion to give the neutral alkoxide intermediate Ru[κ^2 (P,N) PNH] (pcymene) (iPrO) (XI), which transforms into the hydride Ru[κ^2 (P,N) PNH] (p-cymene)H (XII) through β-hydrogen elimination.

From an energetic point of view, Path III is much more disfavored than both Path I and Path

II (see data in Appendix Table 7.13). The less likely steps are the formation of hydride XII and the activation of acetophenone by XII. The first requires in fact the partial decoordination of p-cymene, as the amine is covalently bound to the metal, whereas the activation of acetophenone by a pericyclic transition state is hampered by the sp^2 character of the coordinated amide nitrogen atom, which leads to the formation of a strained and then unfavored RuHCOHN ring.

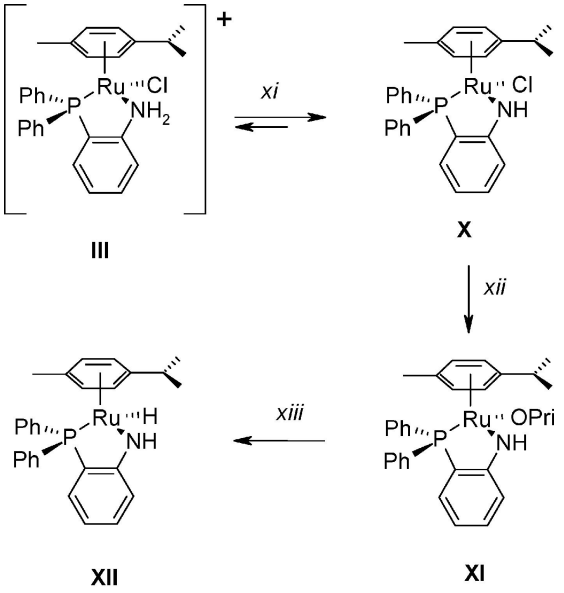


Figure 4.15: Path III - Reaction steps considered in Path III. Conditions: (xi) $iPrO^-$, iPrOH; (xii) $iPrO^-$, Cl^-; (xiii) Me_2CO

4.2.3 *CO* Infrared frequencies correlation with Mulliken charges analysis in Rh catalyst complexes

As mentioned in the Section 3.2.3.1 the zwitterionic mettallate under study in this thesis can be in three different states depending on the pH of solution: neutral (**c1**), cation (**c2**) and bication(**c3**). This kind of chemical system may be seen like a charge-tuned complex, moreover, cause of the particular chemical connections between thioamidic functions and the metal, the charge tuning on the whole molecule reduces its effect on the Rh atom. The total charge centered on the metal, is a mandatory information if we want to have a picture of the activity that these catalysts can display. During the protonation reactions, the electronic distribution

varies dramatically, as evidenced by the C-S and C-N bond distances (153) and by the ^{31}P,^{1}H NMR chemical shift values (246). Moreover, protonation influences the electron density on the Rh(CO) system, as reflected by the infrared CO stretching frequencies (246).

In this contest we apply the computational approach reported in Section 3.2.3.2, describing the electronic distribution in terms of Mulliken charges, frequency analysis and MOs rendering.

We start showing some results about the gas phase calculation of Mulliken charge distribution, looking the Fig. 4.16: it is possible to note that in each of the three cases the most positive charged atoms are the two phosphorous connected by the most negative nitrogen atom close to the metal. This is the first important result because it is in agreement with the strange experimental evidence that locate a dative bond between nitrogen and Rh. Moreover, it is possible to show that the first protonation onto one of the two thioamidic side chain imply a broken symmetry on the optimized geometry. In particular the positive charge additional onto one nitrogen involve an electronic de localization moving from the relative S-Rh bond with consecutive weakening of this. Finally the second protonation leads to an higher symmetry again but with both two S atoms pauperized in electron density. It is reasonable that such electrons come from the metal and the next results will prove this statement.

A complete frequency-population analysis for **c1**, **c2** and **c3** has been performed. The calculations have been made at the BS-0 and BS-I level (see Computational Details in Section 3.2.3.2) in the gas phase and solution. *In vacuo* and at the BS-0 level, the Mullikan charge on the Rh atom varies from -0.338 a.u. to -0.263 a.u. (Table 4.9), showing that the protonation of one thioamidic function induces a 13 % decrease in the electronic population for the mono-protonation and a 22 % decrease for the bi-protonation. Similar results were collected in solution obtaining a decrease of 8 % and 16 % respectively. Such decrease in the calculated population is in agreement with the trend found experimentally for the CO stretching frequencies (246).

As mentioned above a more detailed calculation has been carried out at BS-I level, both in the gas and solution phase, to confirm the BS-0 data. The results are consistent with the above-mentioned trend demonstrating that no errors were induced by usage of pseudo potential to describe the metal (see Appendix Table 7.14). The next step in this study is to try to reproduce not only the above mentioned trend but also experimental quantitative effect of protonation in solution on the CO frequency stretching. In particular we performed standard geometry optimizations on each of the three complexes at the BS-0 level both in the gas and solution phase. On the optimized geometry a full frequency and population analysis has been made with the

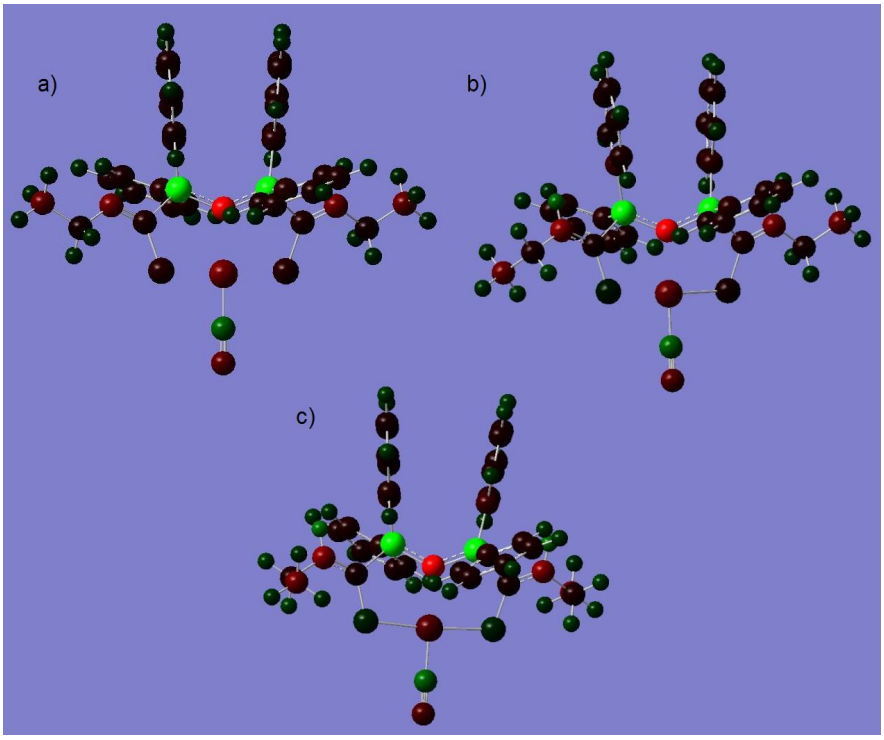
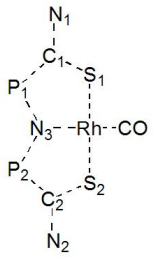


Figure 4.16: Mulliken charge representation - The Mulliken charge is pictured in terms of color scale moving from -0.93 u.a. (red) to +0.93 u.a. (green) for a) neutral, b) monoprotonated and c) diprotonated complex



Gas Phase											
Mol.	Freq.[a]	Rh	S1	S2	N1	N2	N3	C	O	P1	P2
c1	2077.19	-0.338	-0.055	-0.055	-0.381	-0.381	-0.820	0.313	-0.271	0.937	0.937
c2	2113.01	-0.294	0.087	0.0095	-0.409	-0.373	-0.823	0.315	-0.229	0.899	0.922
c3	2144.71	-0.263	0.137	0.134	-0.417	-0.417	-0.825	0.311	-0.188	0.905	0.903
CH_2Cl_2											
Mol.	Freq.[a]	Rh	S1	S2	N1	N2	N3	C	O	P1	P2
c1	2027.01	-0.395	-0.120	-0.116	-0.376	-0.358	-0.835	0.302	-0.280	0.940	0.939
c2	2057.78	-0.361	0.097	-0.074	-0.420	-0.348	-0.836	0.324	-0.264	0.904	0.934
c3	2090.92	-0.331	0.123	0.132	-0.413	-0.402	-0.827	0.337	-0.239	0.912	0.908

Table 4.9: Mulliken populations (a.u.) for selected atoms, in the gas phase and in solution (PCM, CH_2Cl_2) for **c1**, **c2** and **c3**. For compound **c2**, the protonation is on the N1 atom [B3LYP/BS-0//B3LYP/BS-0; BS-0=6-31g** + Lanl2DZ(Rh)]. [a]CO frequencies stretching are in cm^{-1}.

purpose of compare these results with experimental frequency stretching. It is important to point out that such experimental measurements represent the CO bond strength in that complexes in solution and prepared with PF_6 as counter ion. In Fig. 4.17 we report three sets of data which show the linear correlation between electron charge centered on the metal and the frequency stretching of the carbonyl function. The first important observation is the high R^2 obtained in all cases which denotes a strict linear correlation between the two parameters. Although this result encourages to think that experimental and computational linear correlation are in such a way related, it is not possible to establish a good-enough confidence criteria to fix this statement (i.e. failure of the $t - student$ test) .

However, one last point has to be taken into account: as mentioned before the experimental data comes from measurements in solution on complexes prepared with PF_6^- as counter ion (where charged complexes). Moreover, the counter ion represent an interacting distribution of negative charges, both in solid and in solution. It is reasonable to think that a specific interaction may occurs between the protonated thioamidic function and the counter ion. Furthermore, as consequence of the high electron mobility between the metallic center and the thioamidic function, we can expect a stronger or weaker influences of the anion on the CO frequency stretching depending on its coordinative character. We performed several calcula-

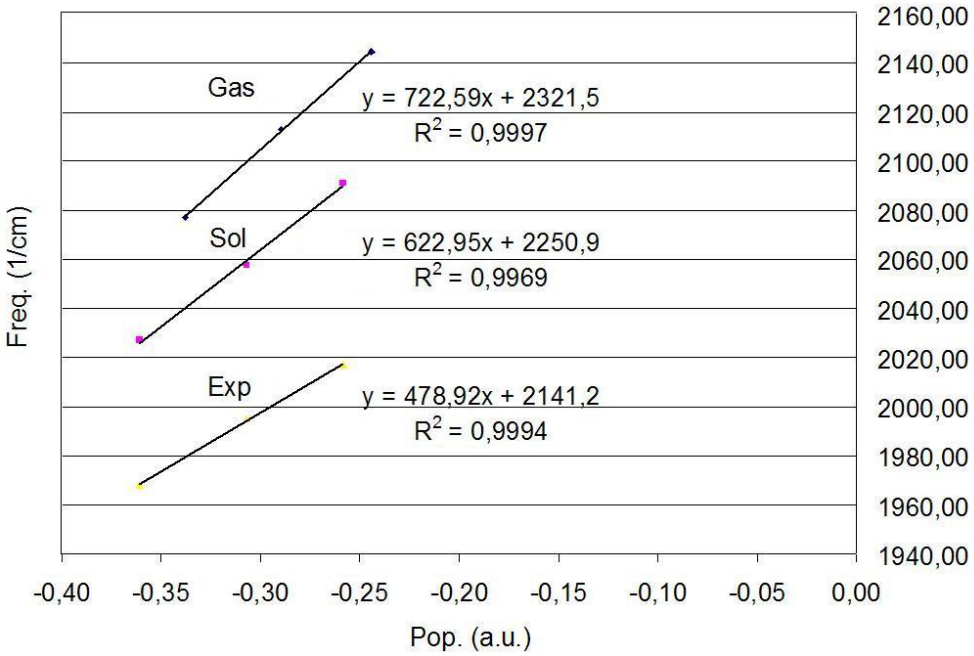


Figure 4.17: Frequency-Mulliken charge correlation - Correlation between the CO frequency stretching and the Mulliken charge computed in gas and in solution. Experimental CO stretching frequencies collected for **c1**, **c2** $\cdot PF_6$ and **c3**$\cdot(PF_6)_2$ are also reported

tions to demonstrate these arguments, with geometry optimization both in gas and in solution of the couple built with the cation and the counter ion involved in an hydrogen bond with the thioamidicfuntion. In particular we focused our attention on PF_6^- and *triflate*, two counter ion with growing coordinative character. In this thesis we just want to show some preliminary results which will be published in detail shortly. In Fig. 4.18 and Table 4.10 it is possible to understand the importance of introducing solvation effects and some specific interactions between mono-, bi- protonated complexes and its counter ion. It is evident that a linear correlation (see R^2 parameter in figure) between CO frequency stretching and the Mulliken charges on the metal is always established. But it is very interesting to note that introducing explicit one- or two- molecules of counter ion such correlation becomes very close to the experimental trend (compare different correlation slopes in table), even better if the frequency correction scaling is applied (i.e. 0.98 (247)).

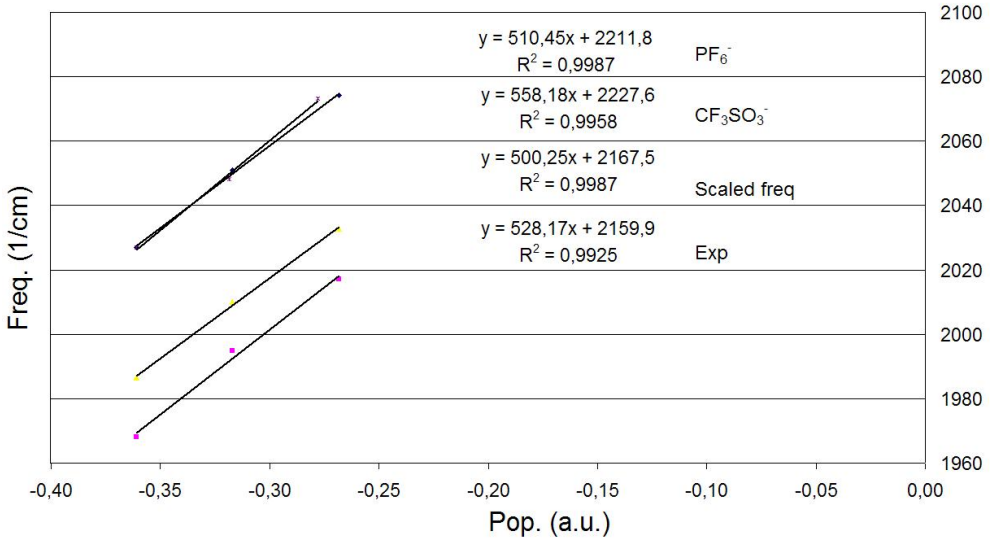


Figure 4.18: Frequency-Mulliken charge correlation - Correlation between the CO frequency stretching and the Mulliken charge computed in solution and in presence of one-, two- molecules of PF_6^- or $CF_3SO_3^-$. Experimental CO stretching frequencies collected for **c1**, **c2** $\cdot PF_6$ and **c3**$\cdot(PF_6)_2$ are also reported and compared with calculated point scaled by 0.98 to correct the B3LYP vibrational frequencies

Mol.	Freq.[a] PF_6^-	Freq.[a] $CF_3SO_3^-$	Freq.[a] PF_6^- scaled	Freq.[a] Exp[b]
c1	2027.01	2027.01	1986.47	1968.00
c2	2051.00	2048.14	2009.98	1995.00
c3	2074.30	2073.27	2032.81	2017.00
Slope[a]	510.45	558.18	500.25	528.17

Table 4.10: Effect of the coordinative character of PF_6^- and $CF_3SO_3^-$ on the CO frequency stretching. Comparison with experimental data end calculation after frequency correction are also reported. [a]CO stretching frequencies stretching are reported in cm^{-1} as well as the correlation slopes. [b] See Ref. in the text

4.3 On the DNA nucleo-base investigation

4.3.1 Selection of the Model Cavity

As anticipated in the previous section, the critical issue for exploring a biological system characterized by several tautomers and equilibrium in solution is to create a model capable of reproducing the experimental pK_a values. In particular, as the pK_as are logarithmic functions of the free energy, a very high accuracy (within 2 kcal/mol) is needed. As has been demonstrated by Goddard's group,(179; 181) the computational level adopted in this work (see Section 3.3.2 is satisfactory for reproducing the gas phase proton affinity of nucleobases. However, there is not currently a standard procedure using the Polarizable Continuum Model (PCM) to predict the solution free energies of the nucleobases with the same accuracy.

The PCM parameter optimized for the test set of DNA nucleobases considered in this study (Fig. 4.20) is the electrostatic scaling factor denoted as which is a real number that is used to increase or decrease the radii of each sphere centered on the atoms (232). The UFF set of atomic radii (see Table 1 in Ref. (233)), with default value of the scaling factor = 1.0 (Gaussian Development Version E05), ensures a good balance between the computational stability and applicability and the reasonable accuracy of the free energy of solution for the neutral solutes (232).

For charged solutes, the scaling factor must be reconsidered. Due to the strong electrostatic interactions, solvent molecules are in general closest to the atoms of charged solutes rather than those of the neutral one.

In a recent study conducted by Camaioni's group, improved quantitative estimates of solvation effects in solution were obtained by decreasing the size of the cavity for cations and anions (248). The representative nucleobase studied to set up the cavities used in this work was guanine and various values of were evaluated by comparing the calculated pK_a value with guanine's experimental first and second pK_a.

Two different values of the scaling factor were adopted: to surround cationic and neutral intermediates a value of 0.91 showed the best results whereas, for anions, a smaller cavity with 0.83 has been used to provide both good accuracy (Fig. 4.19, Appendix Tables 6.1, 6.2 ,6.3, 6.4, 6.5,6.6,6.7,6.8 and Supporting Information of Ref. (249)).

Table 4.11 contains a summary of the relative free energy in the gas and solution phases, the free energy of solution, the calculated and experimental values for the tautomers of guanine and the other nucleobases evaluated during this study. The calculated pK_a values and associated

energy data for each deprotonation of the various guanine tautomers are provided in Appendix Fig. 6.1 and Table 6.9 and Supporting Information of Ref. (249) Scheme S1 and Table S1.

On the basis of previous studies, (181) two cationic, two neutral and three anionic tautomers were considered for the first and second guanine deprotonation (Fig. 4.20). In 4.11, the gas and solution phase energy for each tautomer is expressed relative to the most stable species of each class (e.g. guanine cations relative to the most stable guanine cation).

System	$\Delta G^{a}_{(g)}$ (kcal/mol)	$\Delta G^{*a}_{(aq)}$ (kcal/mol)	$\Delta G^{R'a}_{(sol)}$ (kcal/mol)	pK_a(calc)	pK_a(exp)
Guanine					
G1+	0.0	0.0	-64.8		
G2+	4.3	0.6	-68.9	3.4	3.2-3.3[b]
G3	0.3	1.5	-18.5		
G4	0.0	0.0	-19.0		
G5-	2.8	0.0	-75.4	9.6	9.2-9.6[b]
G6-	0.0	0.3	-72.0		
G7-	0.8	0.7	-73.1		
Adenine					
A1+	1.7	1.5	-60.4		
A2+	8.0	2.8	-65.6		
A3+	0.5	0.0	-60.7		
A4+	0.0	0.3	-59.9	4.2	4.1[c]
A5+	10.1	1.0	-69.5		
A6	0.0	0.0	-9.9		
A7	8.5	2.0	-16.7		
A8	8.5	4.5	-14		
Thymine					
T1	0.0	0.0	-11.1		
T2-	0.0	0.6	-60.3	10.5	9.9[c]
T3-	10.8	0.0	-72.3		
Cytosine					
C1+	0.1	0.0	-65.8		
C2	0.0	0.0	-17.8	4.2	4.4[c]
C3	6.8	3.9	-21.0		
8-oxoG					
8-oxoG1+	7.0	0.0	-82.6	-0.4	-0.1[d]
8-oxoG2	0.0	0.0	-21.0		
8-oxoG3-	0.0	2.0	-69.6	8.0	8.5-8.6[d]
8-oxoG4-	6.8	0.0	-79.1		
Hydantoin					
N1 anion	4.6	6.1	-69.9	11.0	9.16[e]
N3 anion	0.0	0.0	-67.9		

System	$\Delta G^{a}_{(g)}$ (kcal/mol)	$\Delta G^{*a}_{(aq)}$ (kcal/mol)	$\Delta G^{R'a}_{(sol)}$ (kcal/mol)	pK_a(calc)	pK_a(exp)
Phenytoin					
N1 anion	4.2	6.3	-54.1	8.7	8.31[e]
N3 anion	0.0	0.0	-55.8		
Guanidine	0.0	0.0	-62.5	14.5	13.7[f]
N-formyl guanidine					
FG1+	0.0	0.0	-67.0	6.7	not available
FG2+	7.6	3.3	-71.7		
FG3	0.0	0.0	-12.2		
FG8	8.1	0.8	-21.3		
N-acetyl guanidine					
AG1+	0.0	0.0	-61.6	8.5	8.32[f]
AG2+	11.0	7.5	-67.0		
AG3a	0.0	0.0	-9.78		
AG3b	0.1	0.1	-9.78		
AG5	2.0	2.8	-8.54		

Table 4.11: Relative Free Energies (kcal/mol), Calculated and Experimental pK_{a1} Values for Major Tautomers of Guanine, Adenine, Cytosine, Thymine, 8-oxoguanine, Hydantoin, Phenytoin, Guanidine,N-formylguanidine and N-acetylguanidine

Key of Table 4.11 :[a]Energies were calculated using B3LYP density functional theory and the 6-31+G(d,p) and aug-cc-pVTZ basis sets. See equations 3.16,3.20,3.27,3.28. [b]References (181; 248; 250; 251; 252; 253; 254; 255).[c]Reference (256).[d]Reference (257). [e]Reference (258).[f]Reference (259). [g]Reference (260)

The solvation free energy is also reported and clearly shows the effects of this contribution on the relative stability of the three anionic tautomers as the stability of G5- and G6- are reversed between the gas and solution phases. The two pK_a values computed are 3.4 and 9.6 for the first and the second deprotonation respectively, which are in good agreement with the experimental values of 3.2-3.3 (189; 248; 250; 251; 252; 253; 254; 255) and 9.2-9.6 (189; 248; 250; 251; 252; 253; 254; 255).

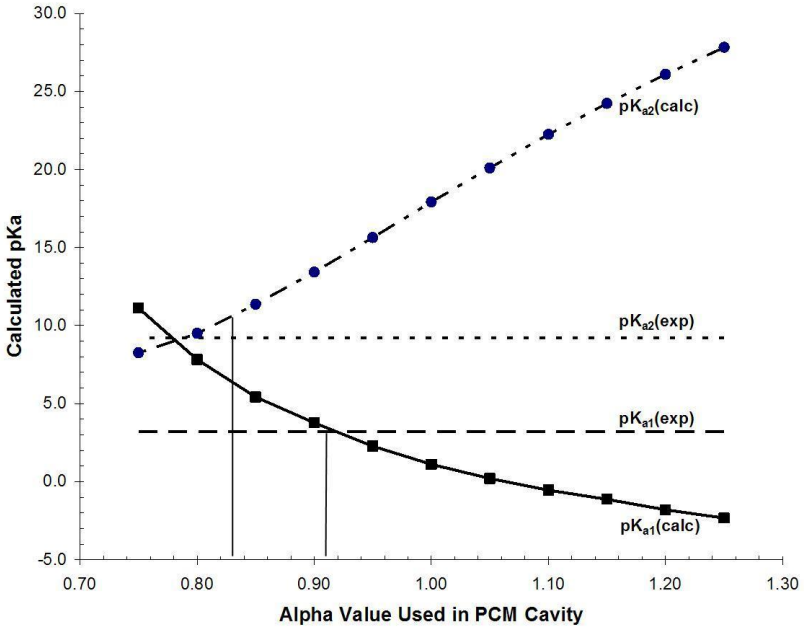


Figure 4.19: Cavity influence - Effect of the UFF alpha value on the calculated pK_{a1} and pK_{a2} values of guanine

4.3.2 Test of the Model Cavity on Other Nucleobases

To be sure that these new cavities could be extended to all the other nucleobases under study (Fig. 4.20), we have performed several test calculations on adenine, cytosine, thymine and 8-oxoG and compared the predicted pK_{a1} to the available experimental data. The data for the most abundant gas and solution phase tautomers are provided in Table 4.11. A complete summary of the data for all tautomers tested for adenine, cytosine, thymine and 8-oxoG is provided in the Appendix Fig. 6.2,6.3,6.4,6.5, in Tables 6.10, 6.11,6.12, 6.13 and Supporting Information Schemes S2-S5 and Tables S2-S5 of Ref. (249).

In the adenine system, the two most stable cationic tautomers are A3+ (protonated at N1) and A4+ (protonated at N3).

In aqueous solution, the two most stable neutral tautomers are predicted to be A6 and A7 which are similar in structure to the lowest energy guanine tautomers, G3 and G4. The global pK_{a1} for adenine deprotonation has been estimated to be 4.2, in very good agreement with the experimental value of 4.1 (248; 251; 252; 253; 254; 255; 261).

The calculations carried out on thymine show essentially only one neutral tautomer, the diketo T1, (Fig. 4.20) and two anions, T2- (deprotonation at N1), and T3- (deprotonation at N3) which are close in energy (Appendix Fig. 6.4 and Table 6.12) and predicted to be present at equilibrium in aqueous solution. All the others enol tautomers are predicted to be either completely absent or at such a low level as to not be relevant to the global pK_{a1} calculation. In this case the calculated pK_{a1} was 10.5 which is in good agreement with the experimental value of 9.9 (248; 251; 252; 253; 254; 255; 261).

Cytosine is a similar system to the pyrimidine base, thymine, but the presence of an amine functional group rather than a carbonyl group at C4 implies the potential existence of several enol imine tautomers (Appendix Fig. 6.3 and Table 6.11). For cytosine, just one cationic and one neutral tautomer are predicted to be relevant in aqueous solution (i.e. C1+ and C2, Table 4.11). Again a good agreement between calculated (4.2) and experimental pK_{a1} (4.4) values has been found.

Finally, in the 8-oxoG system, the data indicate that the presence of the carbonyl group at C-8, instead of -CH, provides a drastic increase of acidity in N3, moving from 2.9 (local pK_{a1} in guanine, see Appendix Fig. 6.1) to -0.4 (local and global first pK_{a1} in 8-oxoG). This is in good agreement with the value found experimentally (0.1). Two different energies contribute to this value: the first one is a large difference between the gas phase basicity which is calculated to be 218.3 kcal/mol for G2+→G4 and 202.8 kcal/mol for 8-oxoG1+→8-oxoG2. These calculations compare favorably with those of Goddard's group (181; 189) which predicted values of 216.18 kcal/mol and 200.8 kcal/mol for the same reactions, respectively, at B3LYP/6-31++G(d,p)//B3LYP/6-31G(d,p).

The experimental value for the gas phase basicity of guanine is 222±2 kcal/mol (262). This energy gap is in part reduced by the contribution of polarization and this is very clear if we consider the solvation free energies of the cationic and neutral intermediates of these two systems. As reported in Table 4.11 for guanine and 8-oxoG, the solvation energies of the neutral tautomers are comparable (-18.5 to 19.0 kcal/mol for guanine versus -21.0 kcal/mol for

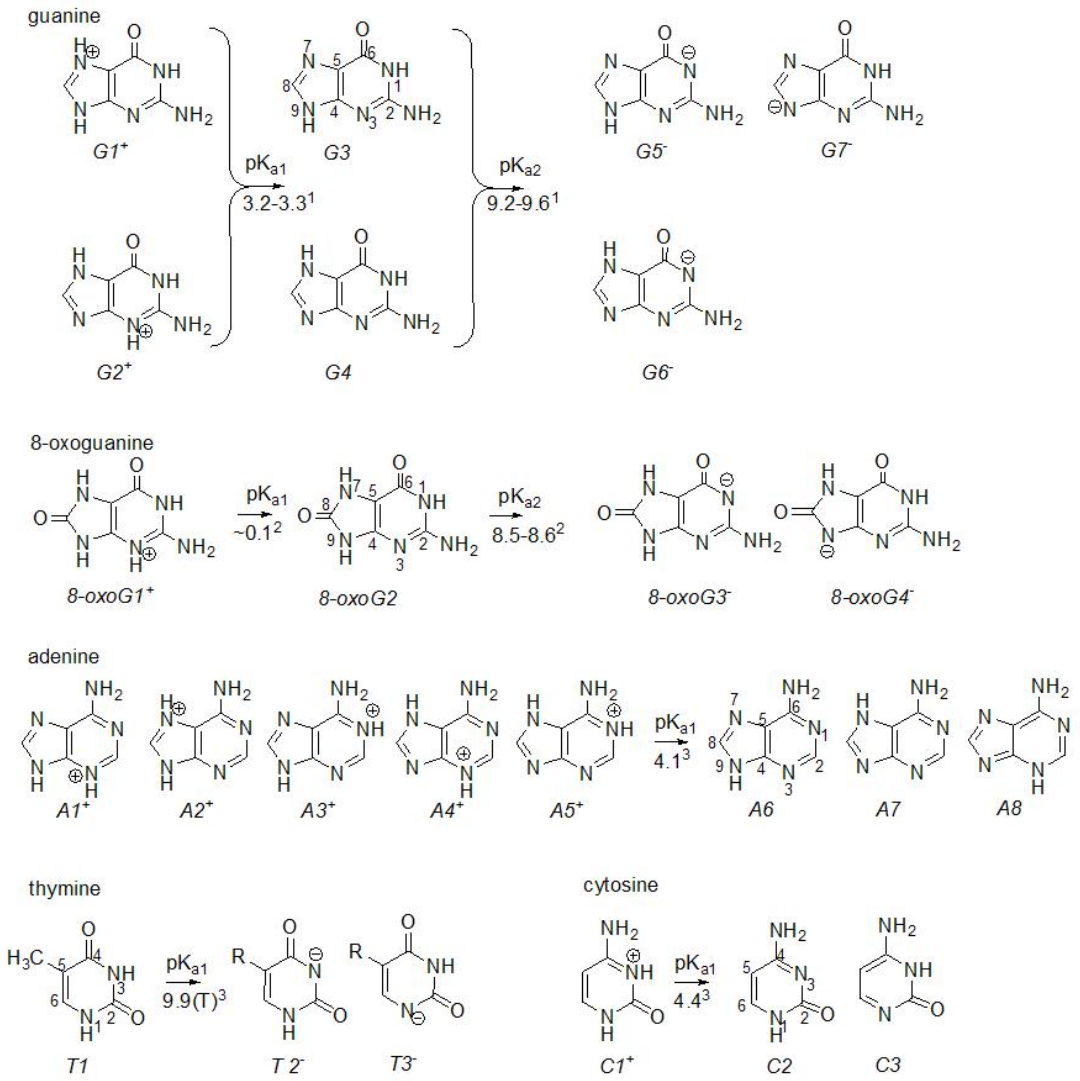


Figure 4.20: Major tautomers of guanine, 8-oxoguanine, adenine, thymine, and cytosine, evaluated for this study - [1]Experimental values for guanine nucleobase (189; 248; 250; 251; 252; 253; 254; 255), [2]Experimental values for 8-oxoguanosine (257), [3]Experimental values for adenine, thymine, uracil and cytosine nucleobases (248; 251; 252; 253; 254; 255; 261). For the adenine modified nucleotide, the measured values for pK_{a1} range from 3.11 to 3.82 (173)

8-oxoG) despite the presence of a more polar group in C8 of the 8-oxoG. In contrast, when the molecule is charged (i.e. cationic tautomers), the free energy of solvation for 8-oxoG is approximately 14-18 kcal/mol greater than for guanine, the energy gap between the 8-oxoG neutral and cationic species decreases leading to a pK_{a1} value near to -0.4.

The second deprotonation of 8-oxoG involves two anionic tautomers and also in this case the calculated pK_{a1} (8.00 units) is close to the experimental value (8.5-8.6 (257)).

4.3.3 Summary of the Computational Method

Based upon the good agreement between the predicted and experimental pK_a values for the five nucleobases, the following steps were taken to calculate the pK_a values for guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp):

1. The geometry of each tautomer was optimized at B3LYP/6-31+G(d,p) and the frequencies were calculated.

2. A gas phase single point calculation was conducted on the gas phase optimized geometry (from Step 1) at B3LYP/aug-cc-pVTZ.

3. The geometry of each tautomer was optimized in aqueous solution at IEF-PCM/B3LYP/6-31+G(d,p) using the gas phase optimized geometry as a starting point. The alpha value for this optimization was the default value of 1.00.

4. A gas phase single point calculation was conducted on the aqueous phase optimized geometry (from Step 3) at B3LYP/aug-cc-pVTZ.

5. The free energy of solvation for the solution phase optimized geometry (from Step 3) was obtained via a single point calculation at IEF-PCM/B3LYP/6-31+G(d,p) using alpha values of 0.91 for the neutral and cationic tautomers and 0.83 for the anions.

6. The free energy in solution of each tautomer and the site-specific or local pK_a values are calculated.

4.3.4 Guanidinohydantoin Results

Several tautomers are possible for guanidinohydantoin for each ionization state and each is expected to contribute to an experimentally observable pK_a in proportion to its population

in solution at 298K. The free energies in the gas and aqueous phases of each of tautomer were calculated using the method described previously (see Section 3.3.2) and are provided in Appendix Tables 7.1 , 7.2, 7.3 and Supporting Information of Ref. (249) Tables S8 and S9. The relative population of each species at 298K was estimated assuming that the tautomers follow the Boltzmann distribution.

4.3.4.1 Tautomers of Neutral Guanidinohydantoin

The nine tautomers of neutral guanidinohydantoin (Gh) considered in this study are shown in Fig. 4.21 and their relative free energies and populations in the gas and aqueous phases are given in Table 4.12. The gas phase electronic energies, sum of the zero point energy and thermal corrections, and free energies of solvation for each species are provided in the Appendix Table 7.4 and Supporting Information of Ref. (249) Table S7.

In the gas phase, the free energy of the neutral tautomers increases in the following order: Gh2 <Gh1 $\ll$ Gh8 <Gh5 <Gh6 <Gh7 <Gh4 <Gh9 <Gh3. The diketo forms of Gh (Gh1, Gh2, and Gh9) are 15 to 26 kcal/mol more stable than the enol tautomers (Gh3-Gh8). This result for the diketo forms of Gh are consistent with NMR, IR, UV, and dipole moment experimental data for hydantoin (Fig. 4.22) which demonstrated that the diketo species is predominant (263; 264; 265; 266; 267).

Earlier semi empirical calculations on hydantoin performed by Kleinpeter et al (263) estimated that the diketo tautomer was 13 to 21.7 kcal/mol (PM3) and 15.6 to 25.2 kcal/mol (AM1) more stable the various imine enol species. Gh2, the N3-C2 imine tautomer, is the most stable species and is 6.4 kcal/mol lower in energy than Gh1, the C2-N1 imine species, due at least in part to the formation of a hydrogen bond between the terminal amino group and the oxygen of the C-5 carbonyl group. Based upon the calculated free energies, the gas phase equilibrium of Gh would consist almost exclusively of the Gh2 tautomer.

In aqueous solution, the free energy of tautomers of neutral Gh increase in a slightly different order from that observed in the gas phase: Gh2 $\approx$ Gh9 <Gh1 $\ll$ Gh8 <Gh5 <Gh6 <Gh7 <Gh4 <Gh3. These data also indicate that the diketo tautomers, Gh1, Gh2 and Gh9, are more stable then the enol tautomers by 14 to 23 kcal/mol. Of the diketo tautomers, Gh2 is predicted to be the most stable species followed by Gh9 (+0.7 kcal/mol), a zwitterion formed by movement of a proton from N7 to N1, and Gh1 (+2.3 kcal/mol). The stability of the zwitterion tautomer containing a protonated guanidine subunit (-N(H)C(NH_2)$_2$) is consistent with the experimentally observed basicity of guanidine (pK_a=13.7) (259) and acidity of hydantoin

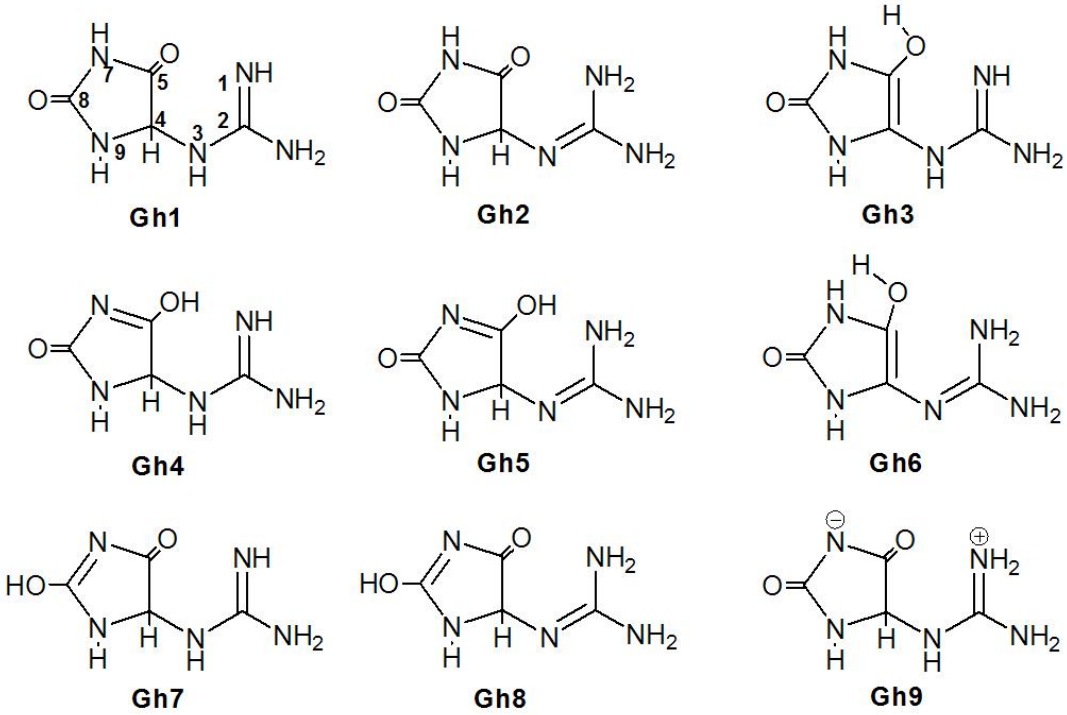


Figure 4.21: Neutral tautomers of guanidinohydantoin evaluated during this study -

	Gh1	Gh2	Gh3	Gh4	Gh5	Gh6	Gh7	Gh8	Gh9
gas phase									
$\Delta G^{0\ a}_{(g)rel}$	6.4	0.0	27.6	24.2	20.9	21.7	23.4	17.5	25.9
Popul.	$1.9x10^{-5}$	1.0	$5.4x10^{-21}$	$1.9x10^{-18}$	$4.3x10^{-16}$	$1.2x10^{-16}$	$6.0x10^{-18}$	$1.4x10^{-13}$	$1.2x10^{-19}$
aq. phase									
$\Delta G^{*\ b}_{(aq)rel}$	2.3	0.0	23.0	18.6	16.3	17.3	na	16.3	0.7
Popul.	0.02	0.76	$9.8x10^{-18}$	$1.9x10^{-14}$	$8.8x10^{-13}$	$1.7x10^{-13}$	na	$7.9x10^{-13}$	0.22

Table 4.12: Relative Free Energies (kcal/mol) and Population of Neutral Tautomers of Guanidinohydantoin (Gh)

Key of Table na = not available. ^{a}Relative energies with respect to $\Delta G^0_{(g)}$ for Gh2. ^{b}Relative energies with respect to $\Delta G^*_{(aq)}$ for Gh2.

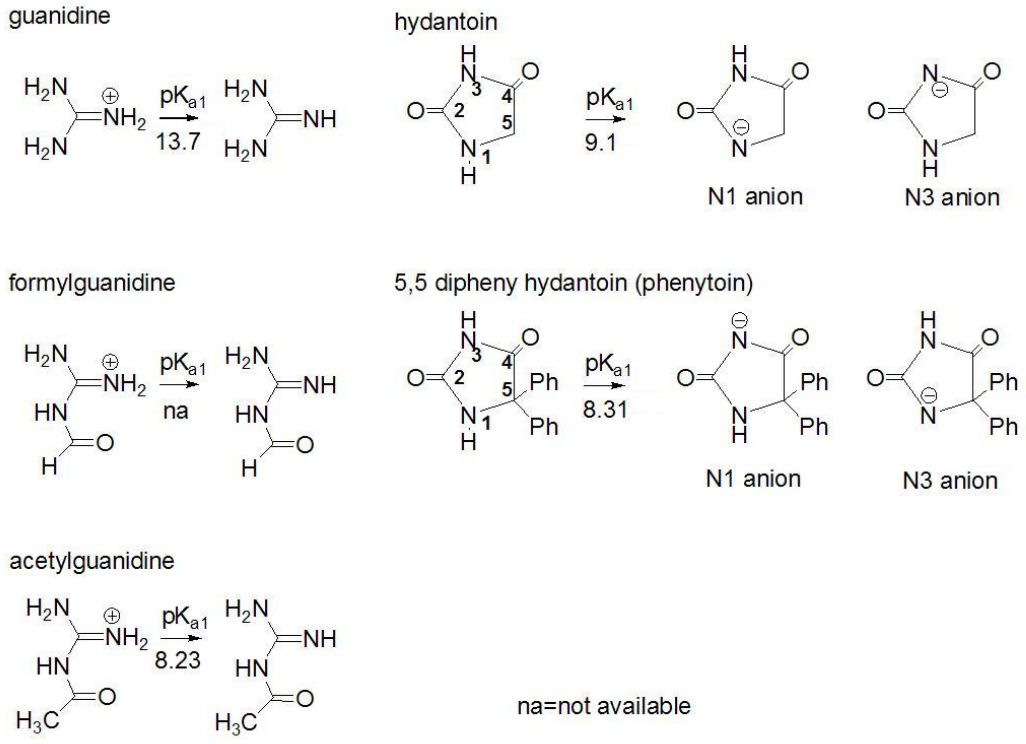


Figure 4.22: Structure and pK_a of guanidine, formylguanidine, acetylguanidine, hydantoin and phenytoin -

(pK_a=9.16) (258) (Fig. 4.22). The equilibrium in aqueous solution is predicted to consist of 76 % Gh2, 22 % Gh9 and 2 % Gh1.

4.3.4.2 Tautomers of Cationic Guanidinohydantoin

Tautomers of cationic guanidinohydantoin considered in this study are shown in Fig. 4.23 and their relative free energies and populations in the gas and aqueous phases are given in Table 4.13. The N-1 site appears to be the major site for protonation in both the gas and solution phases as the data suggest that the equilibrium concentration will consist almost entirely of the $Gh10^+$ tautomer, the diketo hydantoin protonated at N1 of the guanidinyl subunit. This result is consistent with the known acidity and basicity of hydantoin and guanidine. Similar to the pattern observed with the neutral tautomers, the $Gh10^+$ diketo form is significantly more stable than the enol tautomers $Gh13^+$, $Gh14^+$, and $Gh15^+$ in both the gas and solution phases. Protonation of the diketo form of Gh at either the C5 ($Gh11^+$) or C8 ($Gh12^+$) carbonyl oxygen results in a 23.0 to 47.8 kcal/mol increase in free energy in the gas phase and 20.2 to 36.2 kcal/mol increase in aqueous solution.

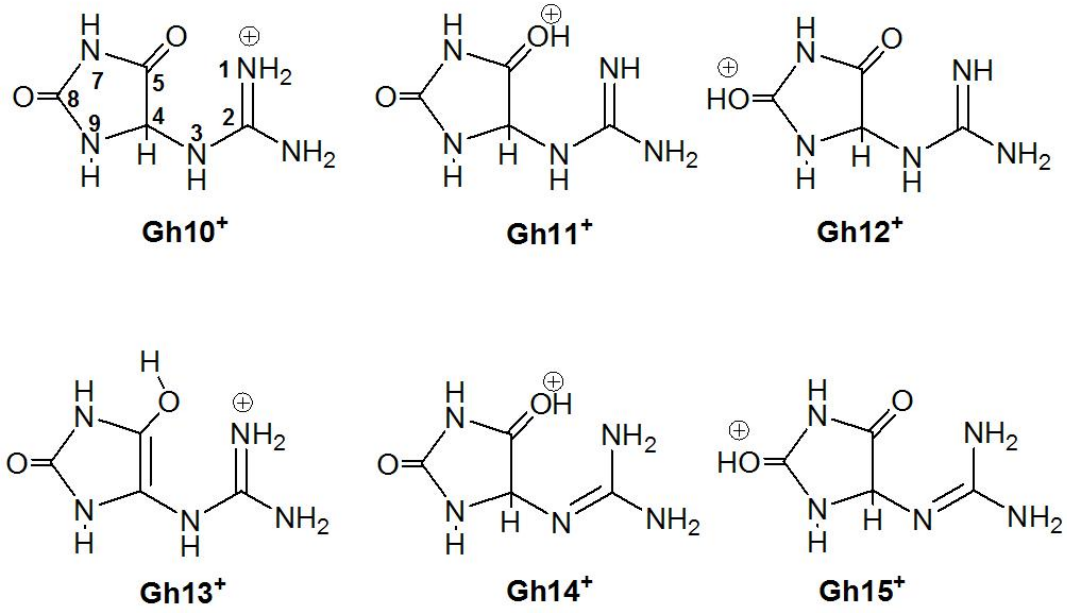


Figure 4.23: Cationic tautomers of guanidinohydantoin evaluated during this study -

	$Gh10^+$	$Gh11^+$	$Gh12^+$	$Gh13^+$	$Gh14^+$	$Gh15^+$
gas phase						
$\Delta G^{0\ a}_{(g)rel}$	0.0	47.8	41.4	23.0	32.1	25.0
Popul.	1.0	$3.7x10^{-29}$	$1.9x10^{-24}$	$5.5x10^{-11}$	$5.5x10^{-11}$	$2.0x10^{-12}$
aq. phase						
$\Delta G^{*\ b}_{(aq)rel}$	0.0	36.2	32.2	20.2	35.5	27.7
Popul.	1.0	$2.8x10^{-27}$	$2.3x10^{-24}$	$1.5x10^{-15}$	$9.38x10^{-27}$	$4.9x10^{-21}$

Table 4.13: Relative Free Energies (kcal/mol) and Population of Cationic Tautomers of Guanidinohydantoin (Gh)

Key of Table ^{a}Relative energies with respect to $\Delta G^0_{(g)}$ for $Gh10^+$. ^{b}Relative energies with respect to $\Delta G^*_{(aq)}$ for $Gh10^+$.

4.3.4.3 Tautomers of Anionic Guanidinohydantoin

Tautomers of anionic guanidinohydantoin considered in this study are shown in Fig. 4.24 and their relative free energies and populations in the gas and aqueous phases are given in Table 4.14. The data indicate that deprotonation at N7H imide position of the Gh diketone tautomers is energetically favored over all other sites as these tautomers, $Gh16^-$ and $Gh20^-$, are at least 4 kcal/mol lower in energy than the next most stable species, $Gh17^-$, a diketo N-9 anion. These results are consistent with NMR data published by Kleinpeter97 for a series of 5,5-disubstituted hydantoins demonstrating that the N3H imide proton (Fig. 4.22) located between the two carbonyl groups was more acidic than the N1H amide proton. Similar to the data for the other Gh ionization states, the enol anions are much higher in energy than the diketo tautomers. As a consequence, in the gas phase, the population of anionic Gh is predicted to be 98 % $Gh16^-$ and 2 % $Gh20^-$. In aqueous solution, the N3-C2 imine species, $Gh20^-$, is 0.9 kcal/mol more stable than $Gh16^-$ and the population ratio is predicted to shift to 18 % $Gh16^-$ and 82 % $Gh20^-$.

4.3.4.4 pK_a of Guanidinohydantoin

Site specific pK_as (pK_a^{ij}) for deprotonation of the various cationic and neutral tautomers of Gh were calculated from the change in free energy for each of the identified reactions provided in Appendix Table 7.3 and Supporting Information of Ref. (249) Table S-9. The global pK_a of Gh was calculated using the Boltzmann weighting method outlined by Goddard (181; 189)

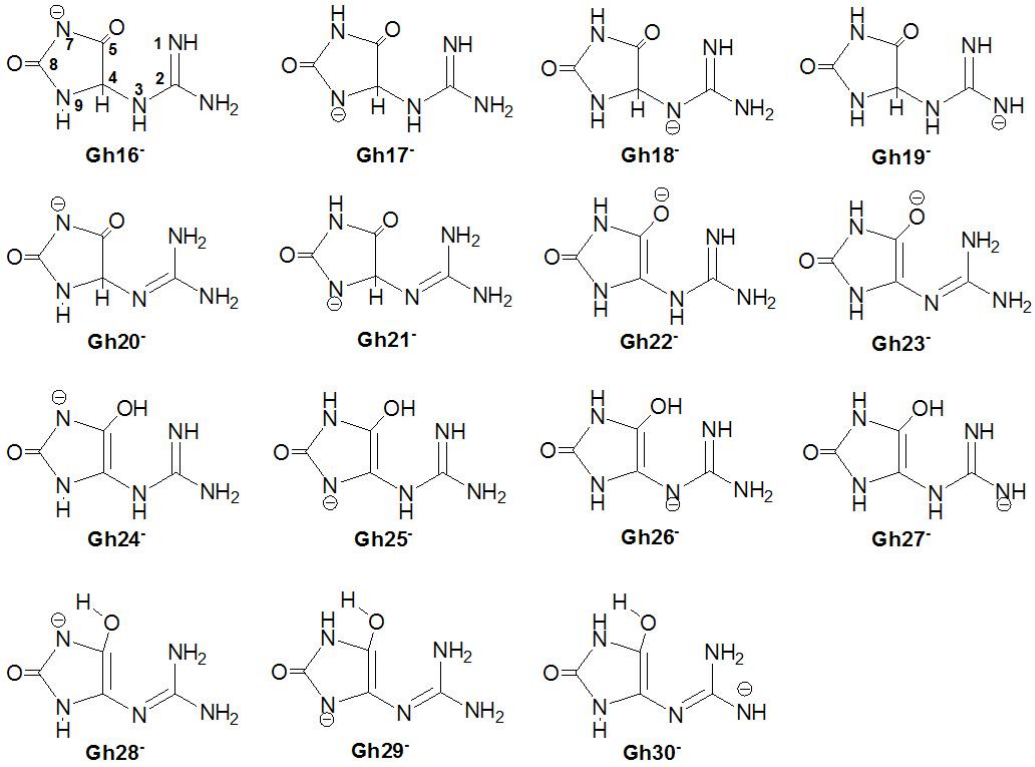


Figure 4.24: Anionic tautomers of guanidinohydantoin evaluated during this study -

	$Gh16^-$	$Gh17^-$	$Gh18^-$	$Gh19^-$	$Gh20^-$	$Gh21^-$	$Gh22^-$	$Gh23^-$
gas phase								
$\Delta G^{0\,a}_{(g)rel}$	0.0	6.4	10.8	26.0	2.2	9.8	19.3	11.3
Popul.	0.98	$1.9x10^{-5}$	$1.1x10^{-8}$	$8.9x10^{-20}$	0.02	$6.3x10^{-8}$	$7.3x10^{-15}$	$5.4x10^{-9}$
aq. phase								
$\Delta G^{*\,b}_{(aq)rel}$	0.9	6.6	14.1	17.7	0.0	8.4	16.0	13.0
Popul.	0.18	$1.2x10^{-5}$	$3.8x10^{-11}$	$8.1x10^{-14}$	0.82	$5.6x10^{-7}$	$1.5x10^{-12}$	$2.3x10^{-10}$
	$Gh24^-$	$Gh25^-$	$Gh26^-$	$Gh27^-$	$Gh28^-$	$Gh29^-$	$Gh30^-$	
gas phase								
$\Delta G^{0\,a}_{(g)rel}$	33.1	30.5	9.7	37.6	27.0	28.3	25.5*	
Popul.	$5.0x10^{-25}$	$4.6x10^{-23}$	$8.2x10^{-8}$	$2.8x10^{-28}$	$1.6x10^{-20}$	$1.6x10^{-21}$	$2.1x10^{-19}$	
aq. phase								
$\Delta G^{*\,b}_{(aq)rel}$	24.8	21.5	12.8	39.0	23.3	25.0	33.0	
Popul.	$5.2x10^{-19}$	$1.4x10^{-16}$	$3.2x10^{-10}$	$2.1x10^{-29}$	$6.3x10^{-18}$	$3.7x10^{-19}$	$5.5x10^{-25}$	

Table 4.14: Relative Free Energies (kcal/mol) and Population of Neutral Tautomers of Guanidinohydantoin (Gh)

Key of Table *The data shown are for the lowest energy rotamer of $Gh30^-$. The relative free energy ranged from 25.5 to 36.0 kcal/mol and from 33.0 to 34.3 kcal/mol in the gas and aqueous phases, respectively. na = not available. ^{a}Relative energies with respect to for $Gh16^-$. ^{b}Relative energies with respect to for $Gh20^-$.

(Equation 3.22) and assumes rapid equilibrium between the various tautomers in aqueous solution. The calculated values for site-specific and global deprotonation of the major tautomers of Gh are summarized in Fig. 4.25.

The predicted pK_{a1} value suggests that deprotonation of the Gh cation will likely occur at a pH of 9.6 at either the N3 position to form Gh2 or at the imide (N7) position to form the Gh9 zwitterionic tautomer. Once formed, the calculated pK_{a2} of 8.2 indicates that these neutral species should almost immediately undergo a second deprotonation to form the two N7 anion species, Gh16- and Gh20^{-}. The data suggests that Gh will behave in a similar manner to its two subunits: guanidine (pK_a = 13.7) and hydantoin (pK_a = 9.16) but is slightly more acidic than either moiety alone.

Using the same computational method discussed previously, the calculated pK_a values for guanidine and hydantoin are predicted to be 14.5 and 11.0, respectively. The increase in acidity of Gh is not unexpected given that some substituted guanidines and hydantoins demonstrate similar behavior. The experimental pK_a values for acetylguanidine and 5,5−diphenylhydantoin (phenytoin) are 8.32 and 8.31, respectively. The method described in this study predicts values of 8.5 for acetyl guanidine and 8.7 for phenytoin (Table 4.11, Appendix Fig. 7.1, Tables 7.5, 7.4 and Supporting Information of Ref. (249) Scheme S7 and Tables S6 and S7).

Experimental values for Gh have not been published to date, however, our results are consistent with SciFinder Scholar's online database for Gh which lists empirically predicted pK_{a1}s and pK_{a2} values of 15.83±0.40 and 12.17±0.70 (calculated using Advanced Chemistry Development software V8.14.) (268). Using the same software, the predicted values for guanidine and hydantoin are 13.27 ± 0.70 and 8.7 ± 0.50, respectively.

4.3.5 Spiroiminodihydantoin Results

Several tautomers are possible for guanidinohydantoin for each ionization state and each is expected to contribute to an experimentally observable pK_a in proportion to its population in solution at 298K. The free energies in the gas and aqueous phases of each of tautomer were calculated using the method described previously and are provided in Appendix Tables 7.6,7.7, 7.8 and Supporting Information of Ref. (249) Tables S10 and S11. The relative population of each species at 298K was estimated assuming that the tautomers follow the Boltzmann distribution.

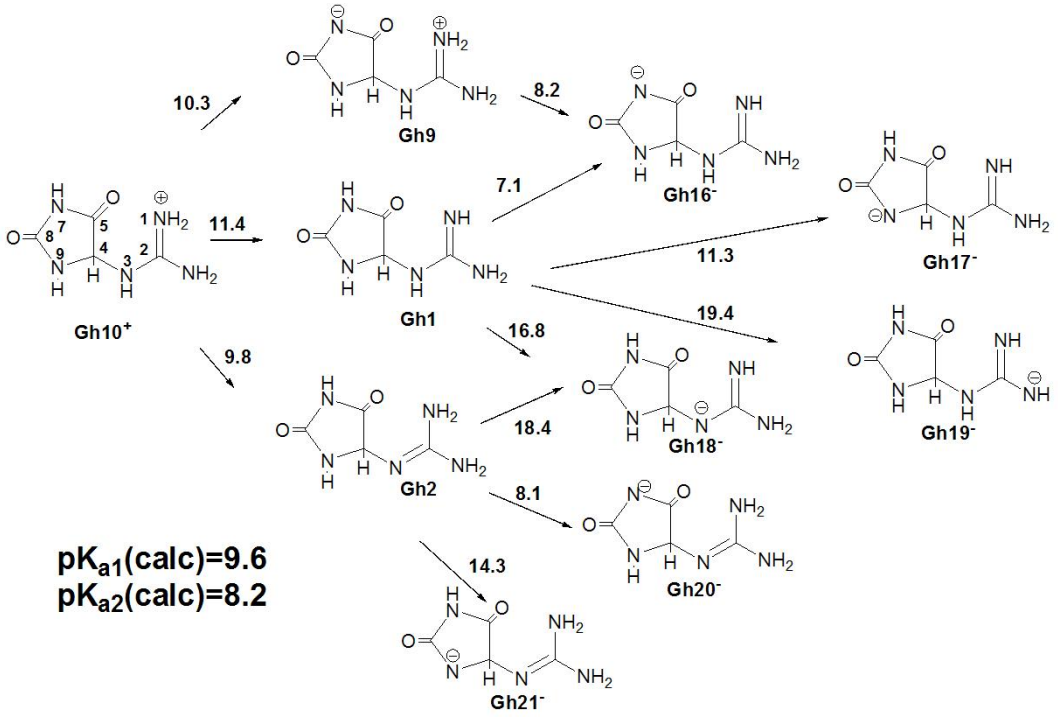


Figure 4.25: Local and global pK_a of major tautomers of guanidinohydantoin evaluated during this study -

4.3.5.1 Neutral Tautomers of Spiroiminodihydantoin

The relative free energies and populations in the gas and aqueous phases of the various neutral tautomers of spiroiminodihydantoin (Sp) considered in this study are given in Table 4.15 and in Tables 7.6, 7.7 7.8 and Tables S10 and S11 of the Supporting Information of Ref. (249); their structures are provided in Fig. 4.26.

In the gas phase, the free energy of each neutral species increases in the following order: Sp1 <Sp3 ≈ Sp2 <Sp4 ≪ Sp8 <Sp15 <Sp11 <Sp14 <Sp5, Sp13 <Sp6, Sp7 <Sp9 <Sp12 <Sp14. Of the 15 tautomers, the four triketo species, Sp1-Sp4 are estimated to be the most stable species with a free energy difference between them of 2 kcal/mol. The remaining 11 tautomers are imine enol species and are at least 10 to 18 kcal/mol higher in energy than Sp1-Sp4, making their population negligible.

In the gas phase, the equilibrium concentration is predicted to be 81 % Sp1 (the C2-N1 imine), 8 % Sp2 (the N3-C2 imine), 8 % Sp3 (a C2-N10 imine) and 3 % Sp4 (a rotamer of Sp3). In aqueous solution, the free energy of the tautomers increases in a slightly different order: Sp1 <Sp2 <Sp3 ≈ Sp4 ≪ Sp10 ≈ Sp9 <Sp11 <Sp8 <Sp6 ≈ Sp5 <Sp7, Sp15 <Sp13 <Sp12 <Sp14. The imine enol species, Sp5-Sp15, are again predicted to be significantly less stable than the four triketo tautomers, Sp1-Sp4. These results are also consistent with the experimental and theoretical data for hydantoin which indicates that the keto species is thermodynamically preferred over the enol tautomers (263; 267). In aqueous solution, the difference in energy between Sp1, the lowest energy species, and Sp2, Sp3 and Sp4 is 3.7, 6.8 and 7.0 kcal/mol respectively. Therefore, at equilibrium, an aqueous solution of Sp will almost exclusively consist of the Sp1 C2-N1 imine tautomer.

4.3.5.2 Cationic Tautomers of Spiroiminodihydantoin

Protonation of Sp can occur at the N1 position and on the oxygen of the carbonyl groups at C5, C6 or C8. The relative free energies and populations in the gas and aqueous phases of the four cationic tautomers of Sp considered in this study are given in Table 4.16; their structures are provided in Fig. 4.27.

Both the gas and solution phase data indicate that $Sp16^+$ is approximately 15 to 26 kcal/mol lower in energy than the other three cations suggesting that Sp is almost exclusively protonated at a guanidinyl nitrogen rather than at one of the carbonyl oxygens.

	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6	Sp7	Sp8
gas phase								
$\Delta G^{0\,a}_{(g)rel}$	0.0	1.4	1.3	2.0	19.5	19.7	19.7	12.2
Popul.	0.81	0.08	0.08	0.03	$4.1x10^{-19}$	$3.0x10^{-15}$	$2.8x10^{-15}$	$8.6x10^{-10}$
aq. phase								
$\Delta G^{*\,b}_{(aq)rel}$	0.0	3.7	6.8	7.0	19.9	19.7	22.1	19.1
Popul.	1.0	$1.8x10^{-3}$	$1.1x10^{-5}$	$7.0x10^{-6}$	$2.6x10^{-15}$	$3.6x10^{-15}$	$6.8x10^{-17}$	$1.1x10^{-14}$
	Sp9	Sp10	Sp11	Sp120	Sp13	Sp14	Sp15	
gas phase								
$\Delta G^{0\,a}_{(g)rel}$	19.9	20.3	18.2	20.0	19.5	19.3	15.6	
Popul.	$2.1x10^{-15}$	$1.1x10^{-15}$	$3.7x10^{-14}$	$1.8x10^{-15}$	$3.9x10^{-15}$	$6.0x10^{-15}$	$3.2x10^{-12}$	
aq. phase								
$\Delta G^{*\,b}_{(aq)rel}$	16.2	15.9	17.5	23.1	22.9	24.2	22.1	
Popul.	$1.3x10^{-12}$	$2.1x10^{-12}$	$1.4x10^{-13}$	$1.2x10^{-18}$	$1.7x10^{-17}$	$1.8x10^{-18}$	$6.3x10^{-17}$	

Table 4.15: Relative Free Energies (kcal/mol) and Population of Neutral Tautomers Spiroiminodihydantoin (Sp)

Key of Table ^{a}Relative energies with respect to for Sp1. ^{b}Relative energies with respect to for Sp1.

	$Sp16^{+}$	$Sp17^{+}$	$Sp18^{+}$	$Sp19^{+}$
gas phase				
$\Delta G^{0\,a}_{(g)rel}$	0.0	21.1	15.7	26.6
Popul.	1.0	$3.7x10^{-16}$	$3.0x10^{-12}$	$3.0x10^{-20}$
aq. phase				
$\Delta G^{*\,b}_{(aq)rel}$	0.0	na	22.3	26.0
Popul.	1.0	na	$4.3x10^{-17}$	$8.2x10^{-206}$

Table 4.16: Relative Free Energies (kcal/mol) and Population of Cationic Tautomers Spiroiminodihydantoin (Sp)

Key of Table na = not available ^{a}Relative energies with respect to for $Sp16^{+}$. ^{b}Relative energies with respect to for $Sp16^{+}$.

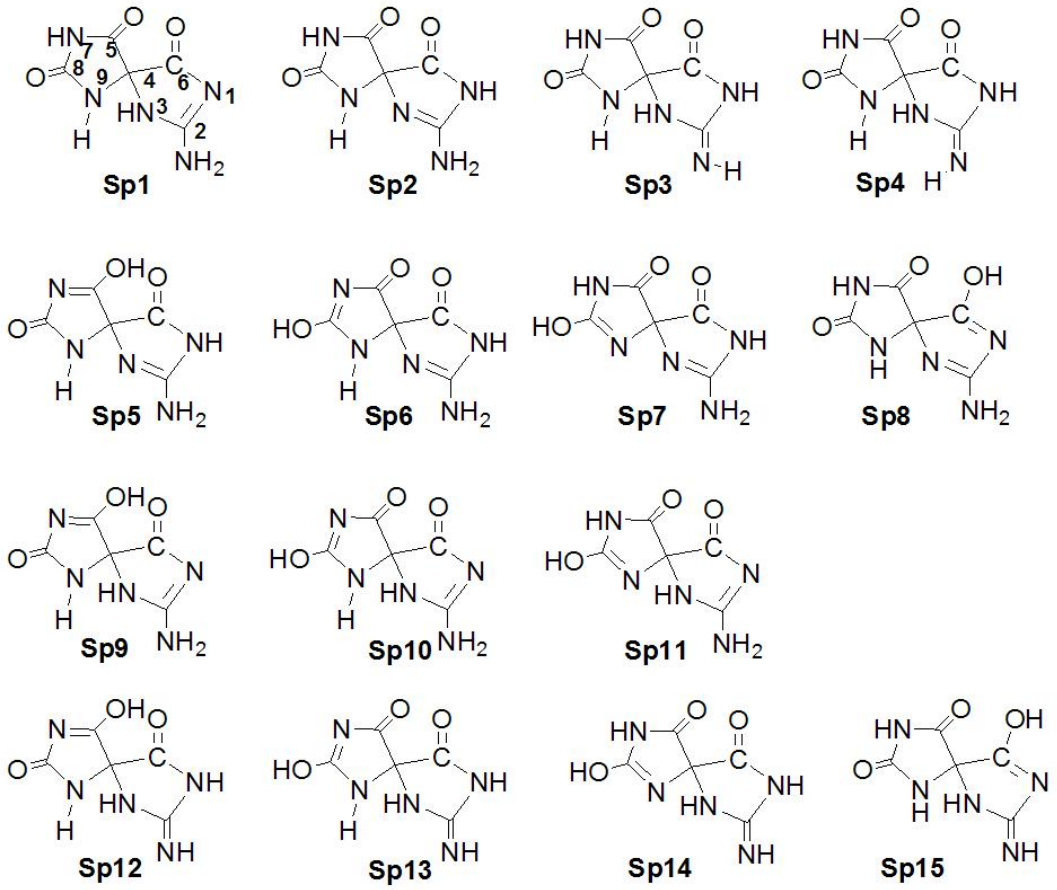


Figure 4.26: Neutral tautomers of spiroiminodihydantoin evaluated during this study -

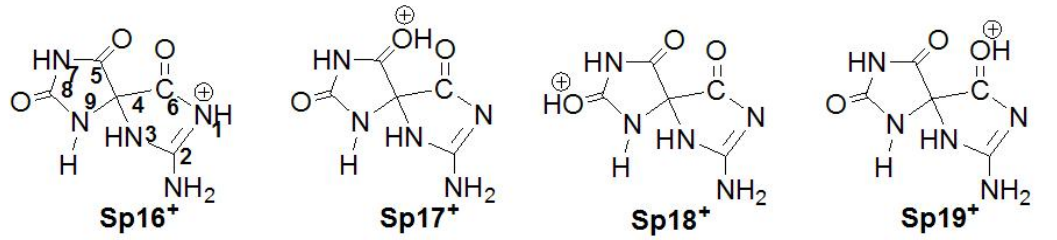


Figure 4.27: Cationic tautomers of spiroiminodihydantoin evaluated during this study -

4.3.5.3 Anionic Tautomers of Spiroiminodihydantoin

Depending on the neutral triketo tautomer, deprotonation of Sp can occur at the N1, N3, N7, N9 positions of the hydantoin rings or from the terminal amino group at C2. For the imine enol tautomers, deprotonation can also occur at the C5, C6 or C8 hydroxyl groups. The relative free energies and populations in the gas and aqueous phases of the 12 anion tautomers of Sp considered in this study are given in Table 4.17; their structures are provided in Fig. 4.28. In the gas phase, the most stable tautomers are $Sp23^-$ (0.0 kcal/mol), $Sp22^-$ (0.86 kcal/mol), $Sp26^-$(1.7 kcal/mol) and $Sp29^-$(3.3 kcal/mol). The free energy of the remaining anionic tautomers increases in the following order: $Sp20^-$ <$Sp31^-$ ≈ $Sp27^-$ <$Sp30^-$ <$Sp28^-$ ≈ $Sp21^-$ ≈ $Sp24^-$ <$Sp25^-$.

In equilibrium in the gas phase, the composition is predicted to comprise 77 % $Sp23^-$, 18 % $Sp22^-$, and 4 % $Sp26^-$. In aqueous solution, the data suggest that the equilibrium composition will consist almost exclusively of the Sp20- N7 anion triketo tautomer. All other tautomers are at least 4.5 kcal/mol higher in energy and will therefore represent less than 0.05 % of the equilibrium population.

	$Sp20^-$	$Sp21^-$	$Sp22^-$	$Sp23^-$	$Sp24^-$	$Sp25^-$
gas phase						
$\Delta G^{0\ a}_{(g)rel}$	5.15	11.8	0.86	0.0	12.0	18.2
Popul.	$1.3x10^{-4}$	$1.6x10^{-9}$	0.18	0.77	$1.2x10^{-9}$	$3.3x10^{-14}$
aq. phase						
$\Delta G^{*\ b}_{(aq)rel}$	0.0	5.3	4.6	7.4	4.5	10.5
Popul.	1.0	$1.4x10^{-4}$	$4.0x10^{-4}$	$3.8x10^{-6}$	$4.7x10^{-4}$	$2.1x10^{-8}$
	$Sp26^-$	$Sp27^-$	$Sp28^-$	$Sp29^-$	$Sp30^-$	$Sp31^-$
gas phase						
$\Delta G^{0\ a}_{(g)rel}$	1.7	6.9	11.6	3.3	8.5	6.4
Popul.	0.04	$7.2x10^{-6}$	$2.4x10^{-94}$	$2.9x10^{-3}$	$4.7x10^{-7}$	$1.5x10^{-5}$
aq. phase						
$\Delta G^{*\ b}_{(aq)rel}$	5.9	11.4	14.5	6.0	11.3	14.4
Popul.	$4.6x10^{-5}$	$4.5x10^{-9}$	$2.2x10^{-11}$	$3.9x10^{-5}$	$5.5x10^{-9}$	$2.8x10^{-11}$

Table 4.17: Relative Free Energies (kcal/mol) and Population of Anionic Tautomers Spiroiminodihydantoin (Sp)

Key of Table ^{a}Relative energies with respect to for $Sp23^-$. ^{b}Relative energies with respect to for $Sp20^-$.

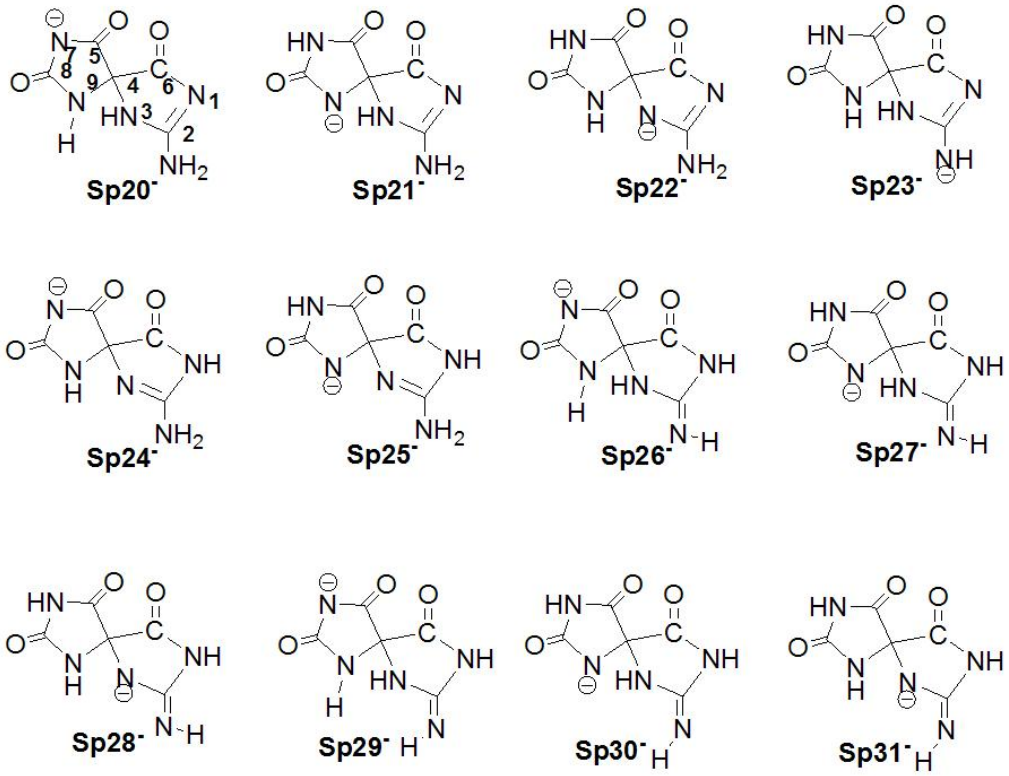


Figure 4.28: Anionic tautomers of spiroiminodihydantoin evaluated during this study -

4.3.5.4 pK_a of Spiroiminodihydantoin

Site specific pK_as (pK_a^{ij}) for deprotonation of the various cationic and neutral tautomers of Sp were calculated from the change in free energy for each of the identified reactions are provided in Appendix Table 7.6 and Supporting Information of Ref. (249) Table S-10.

The global pK_a of Sp was calculated using the Boltzmann weighting method discussed previously (equation 3.22). The calculated values for site-specific and global deprotonation of the major tautomers of Sp are summarized in Fig. 4.29. The predicted pK_{a1} value suggests that deprotonation of the Sp cation will likely occur at a pH of 0.5 and will most likely result from a loss of a proton at the N1 position. This calculated value is in line with both the experimentally observed (~0.1) and calculated (-0.3) values for deprotonation of 8-oxoguanine (179; 181; 189). Experimental data measuring the pK_a of Sp are not available, however SciFinder Scholar's online database for Sp lists empirically predicted values of pK_{a1}=-0.19 and pK_{a2}=7.79 (calculated using Advanced Chemistry Development software V8.14.) (268). Once formed, the calculated pK_{a2} of 4.8 for deprotonation of neutral Sp indicates that this molecule is significantly more acidic than Gh (pK_{a2}=8.2). A close review of the free energy data for the anion tautomers shows that the values for the global pK_{a2} of Sp and Gh are predominately driven by deprotonation at the N7 position of each molecule. These results are consistent with NMR data indicating that the imide proton in both unmodified and 5,5-disubstituted hydantoins (i.e., the N3H proton - Fig. 4.22) is also the most acidic ring proton (263).

For Sp, based upon the Boltzmann weighting, the global pK_{a2} is almost exclusively represented by the deprotonation of Sp1 at N7 leading to formation of Sp20$^-$. For Gh, three reactions contribute to the global pK_{a2}: Gh1→Gh16$^-$ (pK_a=7.1), Gh2→Gh20$^-$ (pK_a=8.1); and Gh9→Gh16$^-$ (pK_a=8.2). The first two reactions are deprotonations at the N7 position; the third is a deprotonation from N1 of the guanidinyl subunit of the zwitterion. The predicted pK_a value of 4.8 for Sp seems surprisingly low for a substituted hydantoin, however, experimental data indicates that the pK_a of hydantoins is very much affected by substituents at N1 and C5 of the ring (263). The experimentally measured pK_a of 5,5-diphenyl hydantoin, sold commercially as phenytoin, is 8.31, almost one full unit lower than the unmodified moiety (269; 270). The predicted value for phenytoin using the method described in this study, is 8.7 (Table 4.11). Modification of phenytoin at the N1 position to yield 1-phenylsulfonyl-5,5-diphenyl hydantoin, reduces the pK_a of the resulting substituted hydantoin to 4.89 (268; 269).

In an effort to examine in detail, the factors contributing to the predicted difference in acidity between Sp and Gh, the contribution of each energy component of the calculation to the site specific pK_a was evaluated and is presented in Table 4.18. Through-space substituent effects were also evaluated and are presented in the same table. Comparing the deprotonation of Sp1 to Gh2 (Gh N3 imine), the major neutral Gh tautomer, the difference in acidity is being driven by the gas phase free energy of the N7 anion relative to its neutral moiety. In the case of Gh1 (C2N1 imine), and the zwitterion, Gh9, the difference in acidity is a consequence of the greater free energy of solvation associated with the Sp anion, $Sp20^-$.

The hydantoin ring containing the N7 proton is the same structure for both Gh and Sp.

One key difference between the two molecules is the presence of a carbonyl group located at the C6 position of Sp. The optimized geometry of Sp orients this carbonyl group directly over the hydantoin ring and it is possible that the lone pair electrons of the oxygen are influencing the acidity of the imine proton (table 4.19). Through-space electronic effects (271) on the acidity of the imine (N7) and amine (N9) protons were evaluated by performing calculations on Sp by replacing the carbonyl at the C6 position with an ethylene group ($C{=}CH_2$) and a CH2 group (Table 4.18).

Deprotonation of four neutral tautomers of $SpC6CH_2$ at the N7 position was evaluated (Appendix Tables 7.9, 7.10 and Supporting Information of Ref. (249) Table S17 and S18). The predicted pK_a of Sp increases significantly as the electron-withdrawing nature of the substituent decreases with calculated values of 6.6 for Sp($C6{=}CH_2$) and 11.9 for Sp($C6H_2$). Visual inspection of the predicted highest occupied molecular orbitals (HOMOs) for Sp1, Sp($C6{=}CH_2$), Sp($C6H_2$) and Gh1 (Table 4.19) indicate that there is significant orbital overlap between the hydantoin ring of Sp1 and the electrons of the C6 carbonyl group. This overlap is not observed for the other species. Through-space electronic effects of the lone-pair may therefore be partly responsible for the predicted increase in acidity of Sp vs Gh.

As indicated by the data in Table 4.17, the predicted relative population of Sp anions in the gas phase and in aqueous solution are significantly different. The difference is, in large part, the result of the approximately 10 kcal/mol larger free energy of solvation of the Sp N7 and N9 anions, $Sp20^-$, $Sp21^-$, $Sp24^-$, and $Sp25^-$, relative to the other anionic tautomers of Sp (Appendix Table 7.3 and Supporting Information of Ref. (249) Table S9). Careful review of the solvation data for these four N7 and N9 anionic tautomers indicates that they also have larger solution phase dipole moments (15.2 to 20.9 debye) relative to the other tautomers (4.6 to 13.7) due to the location of the anion and the lone pairs of the three carbonyl groups. It is

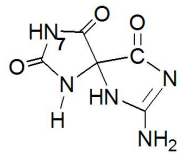
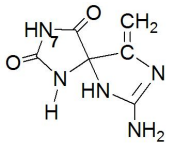
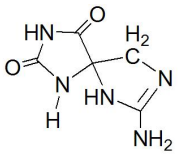
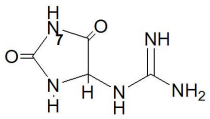
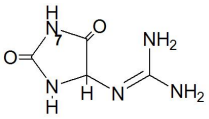
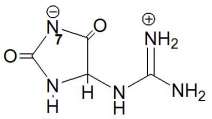
Site-Specific H^+ loss at N7			
Energy Components[a]	Sp1 (C6=O) (1.0)[b]	Sp1 (C6=CH_2)	Sp1 (C6H_2)
$\Delta G_{(g)}$	325.26	328.26	330.27
$\Delta\Delta G^{R'}_{(sol)}$	-56.90	-56.84	-51.37
$\Delta E_{(dis)}$	2.11	1.58	1.28
Site-Specific pK_a	**4.8**	**6.6**	**11.9**
ΔpK_a vs Sp5		+1.8	+7.1
diff $G_{(g)}$		+2.2	+3.7
diff $G_{(sol)}$		+0.0	+4.0
diff $\Delta E_{(dis)}$		-0.4	-0.6
Site-Specific H^+ loss at N7			
Energy Components[a]	Gh1 (0.02)[b]	Gh2 (0.76)[b]	Gh9 (0.22)[b]
$\Delta G_{(g)}$	325.81	334.47	306.36
$\Delta\Delta G^{R'}_{(sol)}$	-54.30	-60.63	-29.7
$\Delta E_{(dis)}$	2.17	1.22	-1.43
Site-Specific pK_a	**7.1**	**8.1**	**8.2**
ΔpK_a vs Sp5	+2.3	+3.3	+3.4
diff $G_{(g)}$	+0.4	+6.7	-13.9
diff $G_{(sol)}$	+1.9	-2.7	+19.9
diff $\Delta E_{(dis)}$	+0.0	-0.7	-2.6

Table 4.18: Acidity of N7 Proton - Comparison of Spiroiminodihydantoin (Sp5) vs Gh7, Gh2 and SpC6CH_2

Key of Table [a]All energies are expressed in kcal/mol and represent the change in energy resulting from deprotonation of the neutral molecule at N7. [b]Boltzmann weighting of neutral molecule

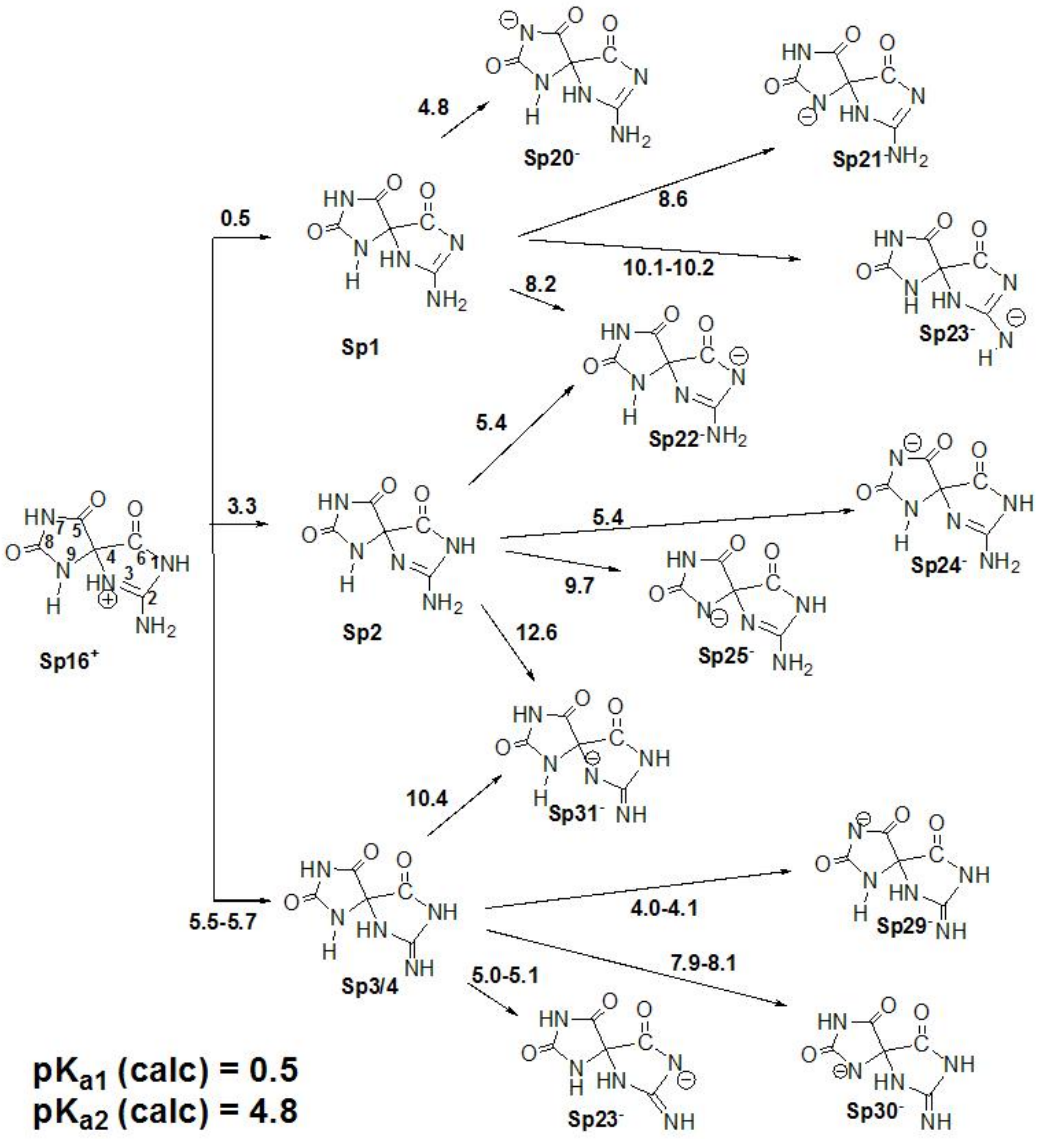


Figure 4.29: Site-specific and global pK_a of major tautomers of spiroiminodihydantoin evaluated during this study -

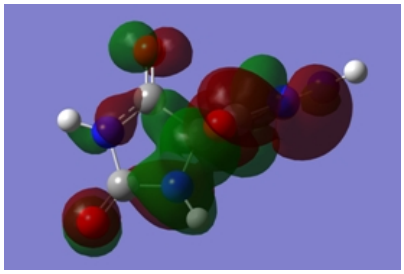
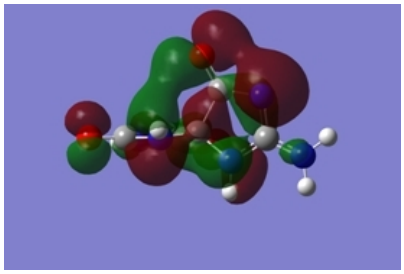
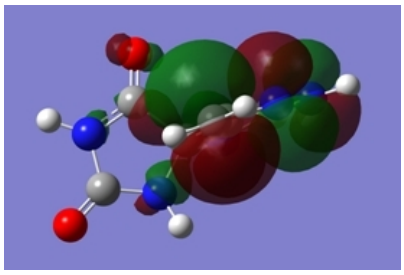
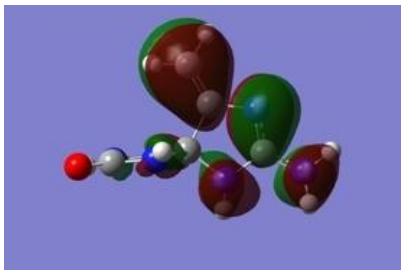
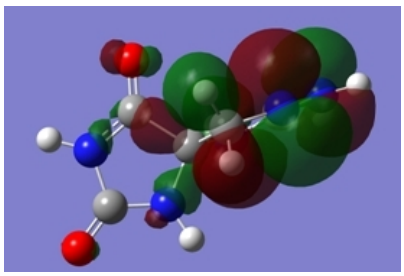
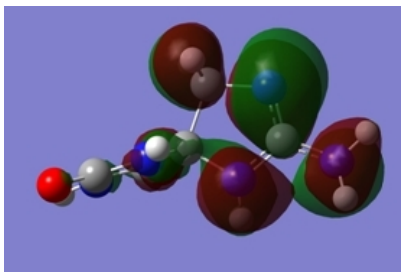
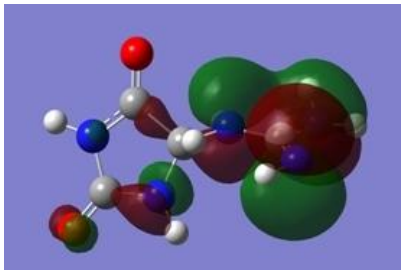
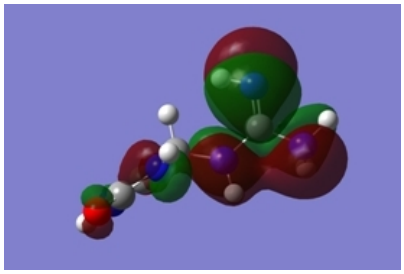
Compound	Front View of Hydantoin Ring	Side View of Hydantoin Ring
Sp1		
Sp1(C6=CH_2)		
Sp1(C6H_2)		
Gh1		

Table 4.19: Comparison of HOMO molecular orbitals for Sp1, Gh1, Sp1(C6=CH_2) and Sp1(C6H_2)

possible that this large polarization of the solute cavity may result in an overestimation of the free energy of solvation by the PCM model and a correspondingly low predicted pK_a value for this molecule.

A similar effect was noted in the calculation of the pK_a of formylguanidine where a value of 6.7 was obtained. One of the tautomers, FG 8 (Appendix Fig. 7.2 and Supporting Information of Ref. (249) Scheme S6), had a calculated free energy of solvation nearly twice that of the other species and a local pK_a value of 5.0. No experimental value for N-formylguanidine was found in the literature, however, SciFinder Scholar's online database lists an empirically predicted value of pK_{a1}=7.75 for this molecule (calculated using Advanced Chemistry Development software V8.14) (268). Based upon the analysis provided above, we believe that the observed pK_{a2} value for Sp is likely to be lower than that of Gh but higher than the 4.8 value predicted by the computational model described in this study.

Calculations conducted with a methylene substituent at the C6 carbon instead of a carbonyl group suggest that the pK_{a2} value for Sp may be in the 6.6 to 7.7 range (Appendix Table 7.10 and Supporting Information of Ref. (249) Table S-18).

5

Conclusions

In this thesis, we have presented a new computational QM method for the study of structural properties (i.e. equilibrium geometry) of molecular systems at very high pressure. The procedure is based on the PCM method, a solvation model largely used to study molecular solutes under normal conditions of pressure. The paper considers two critical items: the definition of the pressure and the elaboration of an analytical code for the calculation of molecular gradients for the regime of high pressure. The reported numerical applications, regarding equilibrium geometries and conformational energies of 1,3-butadiene under high pressure, give an indication of the potentialities of the approach and of the problems to which it may be applied. We address our attention here to further developments of the method. As far as the quantum-mechanical aspect is concerned, the method presented here for SCF approaches (either HF or DFT) can be straightforwardly extended to other QM levels (e.g. MP2, CASSCF) for which the PCM analytical gradients are already available for the electrostatic contribution. Also, the range of the computed molecular observables can be extended. In particular, vibrational frequencies and IR and Raman intensities are observable of primary interest in the study of molecular systems at very high pressure. The analytical gradients presented here could be used to compute the molecular Hessian and the corresponding vibrational frequencies by finite differences. A better approach, however, should use an analytical molecular Hessian under high pressure: extensions in this direction will be presented in the near future. The presented approach can also be extended to the calculation of other second and higher order response properties, e.g. dynamical (hyper) polarizabilities of molecules under high pressure. For these properties, an experimental evidence of the effects of the pressure has been presented and PCM calculations are available at normal pressure (272). These few examples should clearly indicate that the

method proposed in the present work represents a first attempt to improve the theoretical and computational study of molecular systems under high pressure (273).

Moreover, several application in the inorganic chemistry field has been shown and in particular a detailed mechanistic study based on high-resolution MS (ESI) experiments of pre catalytic solutions and DFT/ PCM calculations indicates that the catalytic process is governed by a bifunctional mechanism, analogous to that proposed for bis(amine) and amino alcohol ligands.(146; 274; 275; 276) The two active catalytic intermediates are the 16e- cation Ru[(κ^2P,N)PNH](p-cymene)$^+$ (IV) and the cationic hydride Ru[(κ^2P,N)PNH$_2$](p-cymene)H$_+$ (V). DFT/PCM calculations show that the oxidation of iPrOH and reduction of acetophenone occur by concerted hydride and proton transfers by pericyclic transition states. Although DFT calculations predict the possibility that V could also be formed through β-hydrogen elimination from an alkoxide precursor (Fig. 4.8), experimental evidence supporting this proposal was not collected. The importance of the NH_2 functionality on catalysis is underlined by the low activity of 2a under the same catalytic conditions. Although the hydride Ru[(κ^2P,N)PNMe$_2$](p-cymene)H quickly forms in basic iPrOH, the activation of acetophenone, likely through detachment of the amine functionality, is hampered. Finally, a mechanism operating through neutral organometallic intermediates of the type Ru[(κ^2P,N)PNH](pcymene)(X) has been ruled out on the basis of DFT results.

Finally an efficient computational method has been identified which uses B3LYP density functional theory, IEF-PCM solvation modeling with a modified UFF cavity and Boltzmann weighting of tautomers to predict the site-specific and global pK_a of DNA nucleobases and their oxidation products. The method is shown to be capable of predicting the global pK_a of the DNA nucleobases guanine, 8-oxoG, adenine, cytosine and thymine to within 0.6 pK_a units of their experimental value. Predictions of the experimental values of phenytoin and N-acetylguanidine were within 0.4 pK_a units. The method works less well for smaller molecules such as guanidine and hydantoin where the calculated values are higher by 0.8 and 1.9 pK_a units, respectively. The method has been used to evaluate the acidity of Gh and Sp, two highly mutagenic guanine oxidation products. The trend observed for the pK_a values of Gh (9.6 and 8.2) is consistent with the experimentally observed values for guanidine cation (13.7) and hydantoin (9.16). Molecular orbital predictions of the HOMO indicate that there is very little interaction between the hydantoin and guanidine subunits of Gh. The pK_{a1}(calc) value for deprotonation of Sp cation (Sp$^+$ $\rightarrow$Sp) is very close to the experimentally observed pK_{a1} for 8-oxoG and is consistent with the similarity in their structures. The data suggest that the imide

(N7) proton in Sp may be more acidic than that in Gh, possibly due to the presence of the through-space electronic effects of the carbonyl group located at C6. This difference in the acidity of Gh and Sp may be an indication of their potential toxicity and mutagenicity in vivo and remains a fertile area for experimental study.

6

Appendix

REACTION LIST

PATH I

1. **III** + i-PrO$^-$ + i-PrOH + PhC(O)CH_3 + Cl^-
2. **X** + 2i-PrOH + PhC(O)CH_3 + Cl^-
3. **IV** + 2i-PrOH + PhC(O)CH_3 + 2Cl^-
4. **TS2** + i-PrOH + PhC(O)CH_3 + 2Cl^-
5. **V** + (CH3)2CO + i-PrOH + PhC(O)CH_3 + 2Cl^-
6. **TS3(Si)** + (CH3)2CO + i-PrOH + 2Cl^-
7. **TS3(Re)** + (CH3)2CO + i-PrOH + 2Cl^-
8.) **IV** + (CH3)2CO + i-PrOH + PhCH(OH)CH_3 + 2Cl^-

PATH II

1. **III** + i-PrO- + i-PrOH + PhC(O)CH_3 + Cl^-
2. **IX** + i-PrOH + PhC(O)CH_3 + 2Cl^-
3. **TS1** + i-PrOH + PhC(O)CH_3 + 2Cl^-
4. **V** + (CH3)2CO + i-PrOH + PhC(O)CH_3 + 2Cl^-

5. **TS3(Si)** + (CH3)2CO + i-PrOH + $2Cl^-$
6. **TS3(Re)** + (CH3)2CO + i-PrOH + $2Cl^-$
7. **IV** + (CH3)2CO + i-PrOH + $PhCH(OH)CH_3$ + $2Cl^-$
8. **TS2** + (CH3)2CO + $PhCH(OH)CH_3$ + $2Cl^-$

List of reaction for the two possible paths into the hydrogen transfer

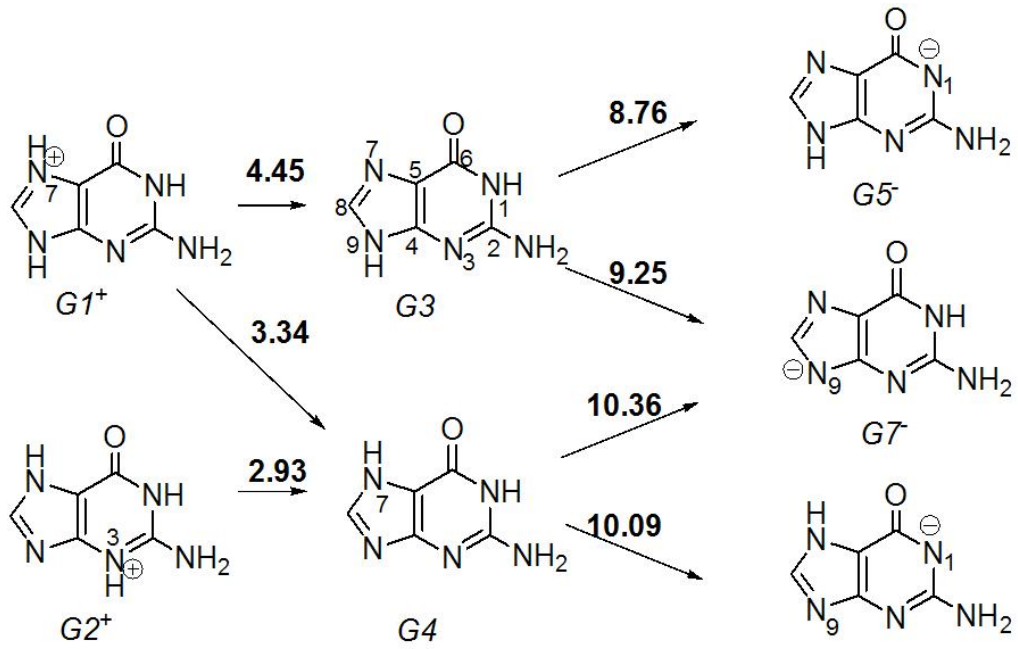


Figure 6.1: Appendix - Guanine tautomers and their local pK_a values

Tautomer	$G5^-$					
α=	$\Delta G^{0\,R'}_{(sol)}$ (Kcal/mol)	$\Delta G^{0\,R}_{(g)}$ (Kcal/mol)	$\Delta E^{0}_{(dis)}$ (Kcal/mol)	$G^{0}_{(aq)}$ (Kcal/mol)	G3 $\rightarrow G5^-$ pKa local	Global pKa
1.25	-51.25	-340198.25	1.31	-340248.19	26.98	27.84
1.20	-53.69			-340250.63	25.19	26.09
1.15	-56.31			-340253.25	23.27	24.24
1.10	-59.15			-340256.09	21.19	22.25
1.05	-62.13			-340259.07	19.01	20.10
1.00	-65.19			-340262.13	16.76	17.92
0.95	-68.39			-340265.33	14.42	15.63
0.90	-71.45			-340268.39	12.18	13.42
0.85	-74.33			-340271.27	10.07	11.37
0.80	-76.89			-340273.83	8.19	9.51
0.75	-78.65			-340275.59	6.90	8.26
Tautomer	$G6^-$					
α=	$\Delta G^{0\,R'}_{(sol)}$ (Kcal/mol)	$\Delta G^{0\,R}_{(g)}$ (Kcal/mol)	$\Delta E^{0}_{(dis)}$ (Kcal/mol)	$G^{0}_{(aq)}$ (Kcal/mol)	G3 $\rightarrow G6^-$ pKa local	Global pKa
1.25	-48.58	-340201.08	1.02	-340248.64	27.80	27.42
1.20	-50.96			-340251.02	26.05	25.67
1.15	-53.51			-340253.57	24.18	23.81
1.10	-56.23			-340256.29	22.19	21.82
1.05	-59.14			-340259.20	20.06	19.68
1.00	-62.11			-340262.17	17.88	17.49
0.95	-65.21			-340265.27	15.61	15.21
0.90	-68.19			-340268.25	13.42	12.99
0.85	-70.92			-340270.98	11.42	10.93
0.80	-73.44			-340273.50	9.58	9.07
0.75	-75.13			-340275.19	8.34	7.82

Table 6.1: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with Various Alpha Values for UFF Cavity

Tautomer	$G7^-$							
α=	$\Delta G^{0\,R'}_{(sol)}$	$\Delta G^{0\,R}_{(g)}$	$\Delta E^{0}_{(dis)}$	$G^{0}_{(aq)}$	G3 →G7$^-$	G4 →G7$^-$	G3 →G7$^-$	G4 →G7$^-$
	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	pKa local	pKa local	Global pKa	Global pKa
1.25	-50.09	-340200.25	1.66	-340248.67	26.63	27.77	27.42	27.42
1.20	-52.42			-340251.00	24.92	26.06	25.69	25.68
1.15	-54.94			-340253.52	23.07	24.21	23.80	23.80
1.10	-57.55			-340256.13	21.16	22.30	21.83	21.82
1.05	-60.43			-340259.01	19.05	20.19	19.67	19.67
1.00	-63.35			-340261.93	16.91	18.05	17.48	17.48
0.95	-66.36			-340264.94	14.70	15.84	15.21	15.21
0.90	-69.39			-340267.97	12.48	13.62	12.98	12.98
0.85	-72.14			-340270.72	10.47	11.61	10.92	10.92
0.80	-74.62			-340273.20	8.65	9.79	9.08	9.08
0.75	-76.18			-340274.76	7.51	8.65	7.83	7.83
Tautomer	$G1^+$							
α=	$\Delta G^{0\,R'}_{(sol)}$	$\Delta G^{0\,R}_{(g)}$	$\Delta E^{0}_{(dis)}$	$G^{0}_{(aq)}$	G1 →G3$^+$	Global pKa		
	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	pKa local			
1.25	-42.24	-340766.70	0.59	-340808.35	-1.80	-2.33		
1.20	-44.38			-340810.49	-1.19	-1.80		
1.15	-46.84			-340812.95	-0.51	-1.13		
1.10	-49.42			-340815.53	0.15	-0.55		
1.05	-52.59			-340818.70	0.97	0.20		
1.00	-56.32			-340822.43	1.96	1.11		
0.95	-60.74			-340826.85	3.20	2.27		
0.90	-65.93			-340832.04	4.77	3.76		
0.85	-71.66			-340837.77	6.36	5.42		
0.80	-79.11			-340845.22	8.77	7.83		
0.75	-88.22			-340854.33	12.11	11.11		

Table 6.2: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with Various Alpha Values for UFF Cavity

Tautomer	$G2^+$					
α=	$\Delta G^{0\,R'}_{(sol)}$	$\Delta G^{0\,R}_{(g)}$	$\Delta E^{0}_{(dis)}$	$G^{0}_{(aq)}$	$G2^+ \rightarrow$G4	Global pKa
	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	pKa local	
1.25	-45.42	-340762.35	0.92	-340806.85	-3.32	-2.32
1.20	-47.72			-340809.15	-2.69	-1.78
1.15	-50.37			-340811.80	-1.95	-1.22
1.10	-52.97			-340814.40	-1.35	-0.61
1.05	-56.36			-340817.79	-0.47	0.23
1.00	-60.23			-340821.66	0.52	1.14
0.95	-64.75			-340826.18	1.70	2.28
0.90	-70.14			-340831.57	3.29	3.76
0.85	-76.48			-340837.91	5.20	5.43
0.80	-84.12			-340845.55	7.65	7.83
0.75	-93.54			-340854.97	11.00	11.12
Tautomer	G3					
α=	$\Delta G^{0\,R'}_{(sol)}$	$\Delta G^{0\,R}_{(g)}$	$\Delta E^{0}_{(dis)}$	$G^{0}_{(aq)}$		
	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)		
1.25	-4.46	-340537.11	1.03	-340540.54		
1.20	-5.77			-340541.85		
1.15	-7.30			-340543.38		
1.10	-8.99			-340545.07		
1.05	-11.04			-340547.12		
1.00	-13.41			-340549.49		
0.95	-16.14			-340552.22		
0.90	-19.19			-340555.27		
0.85	-22.75			-340558.83		
0.80	-26.91			-340562.99		
0.75	-31.46			-340567.54		

Table 6.3: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with Various Alpha Values for UFF Cavity

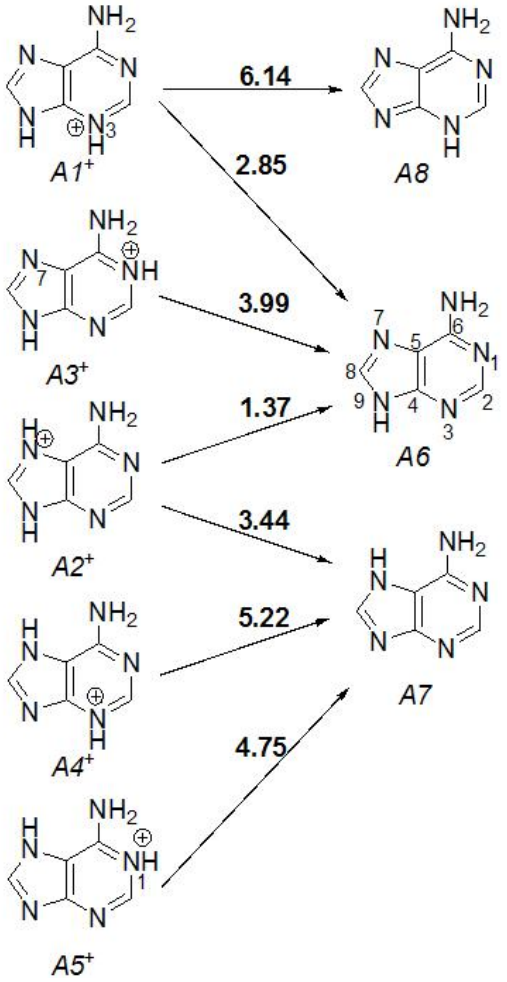


Figure 6.2: Appendix - Adenine tautomers and their local pK_a values

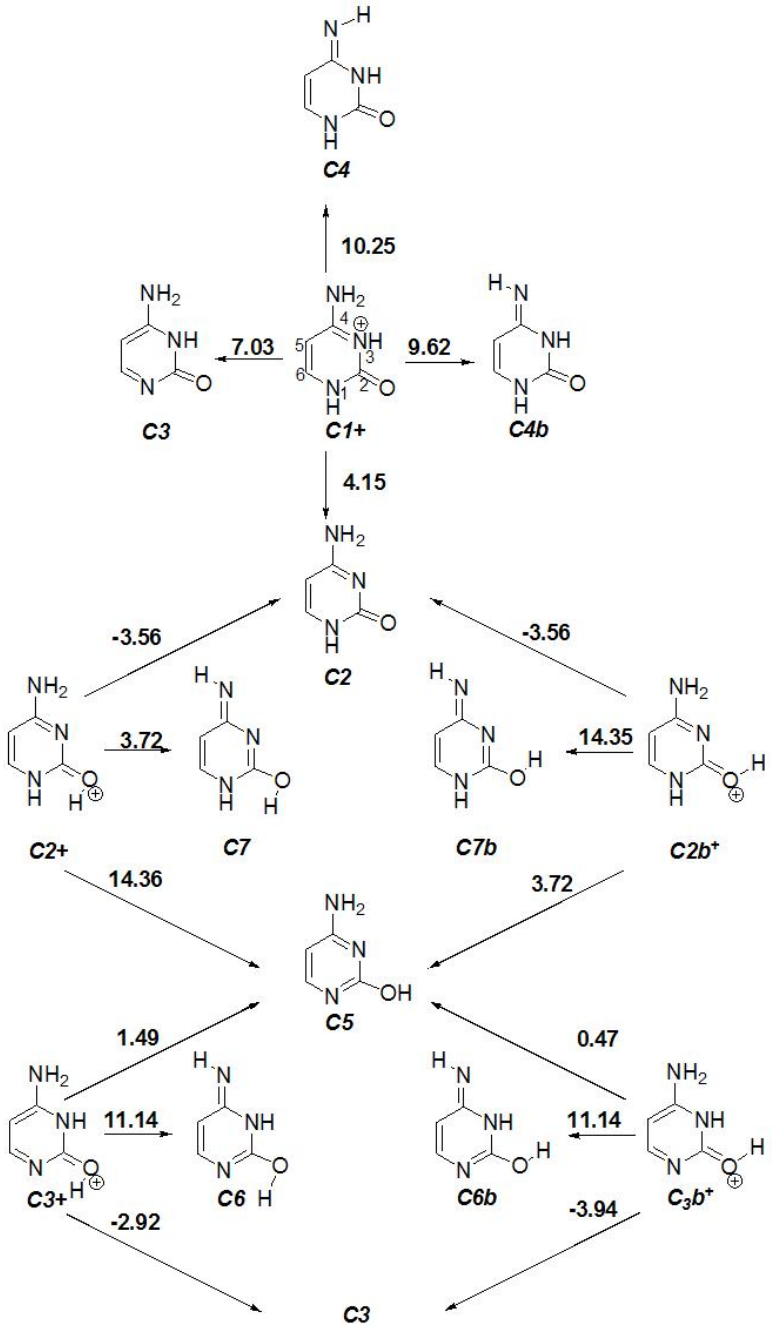


Figure 6.3: Appendix - Cytosine tautomers and their local pK_a values

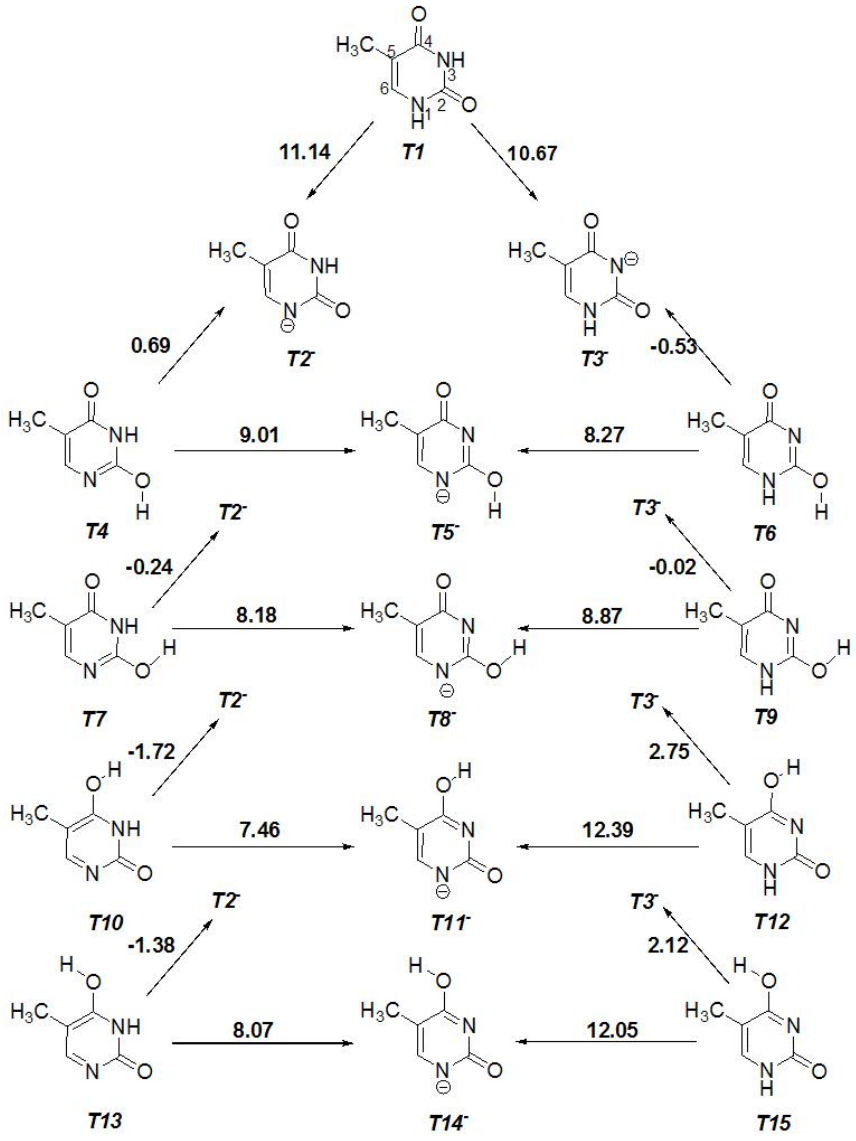
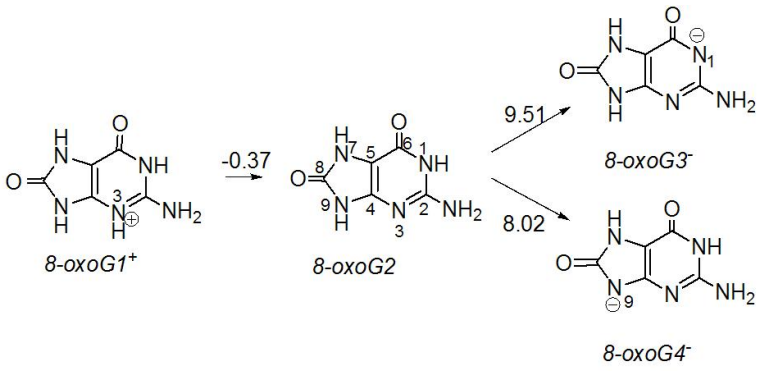
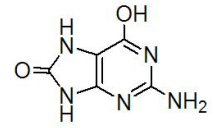
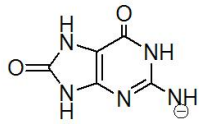


Figure 6.4: Appendix - Thymine tautomers and their local pK_a values



Other tautomers evaluated at UFF, α=0.91





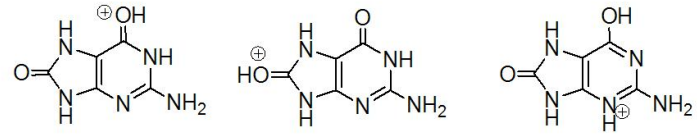


Figure 6.5: Appendix - 8-oxoguanine tautomers and their local pK_a values

Tautomer α=	G4 $\Delta G^{0\,R'}_{(sol)}$ (Kcal/mol)	$\Delta G^{0\,R}_{(g)}$ (Kcal/mol)	$\Delta E^{0}_{(dis)}$ (Kcal/mol)	$G^{0}_{(aq)}$ (Kcal/mol)	$G1^{+} \rightarrow$G4 pKa local	Global pKa
1.25	-4.08	-340537.75	0.71	-340541.12	-2.23	-2.33
1.20	-5.52			-340542.56	-1.71	-1.78
1.15	-7.15			-340544.19	-1.10	-1.13
1.10	-8.93			-340545.97	-0.52	-0.53
1.05	-11.13			-340548.17	0.19	0.21
1.00	-13.65			-340550.69	1.08	1.13
0.95	-16.55			-340553.59	2.19	2.28
0.90	-19.78			-340556.82	3.63	3.77
0.85	-23.51			-340560.55	5.10	5.43
0.80	-27.80			-340564.84	7.41	7.83
0.75	-32.65			-340569.69	10.53	11.12

Table 6.4: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with Various Alpha Values for UFF Cavity

Tautomer	$E^0_{(g)}(R)$[1]	ZPE[2]+ $\Delta G_{0\rightarrow 298}$	$\Delta G^0_{(g)}(R)$	$E^0_{(g)}(R')$[1]	$\Delta E^0_{(dis)}$	$\Delta G^{0\,R'}_{(sol)}$[3]	$G^0_{(aq)}$	Rel $G^0_{(g)}$	Boltzmann	Rel $G^0_{(aq)}$
	(hartrees)	(hartrees)	(Kcal/mol)	(hartrees)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	Pop.(g)	(Kcal/mol)
G1+	-543.1430	0.0967	-340766.70	-543.1421	0.59	-66.71	-340832.81	0.00	9.96E-01	0.00
G2+	-543.1353	0.0959	-340762.35	-543.1338	0.92	-71.32	-340832.75	4.35	0.64E-03	0.06
G3	-542.7640	0.0836	-340537.11	-542.7624	1.03	-19.87	-340555.95	0.65	2.40E-01	1.62
G4	-542.7652	0.0837	-340537.75	-542.7640	0.71	-20.52	-340557.57	0.00	7.19E-01	0.00
G5-	-542.2105	0.0701	-340198.25	-542.2084	1.31	-72.02	-340268.96	2.82	0.64E-02	0.00
G6-	-542.2154	0.0705	-340201.08	-542.2138	1.02	-68.78	-340268.84	0.00	7.50E-01	0.13
G7-	-542.2141	0.0705	-340200.25	-542.2114	1.66	-69.96	-340268.54	0.83	1.85E-01	0.42
G17–	-541.5191	0.0575	-339772.26	-541.5171	1.22	-202.37	-339973.41	0.00	1.00	0.00
									Pop.(aq)	
G1+									5.25E-01	
G2+									4.75E-01	
G3									6.10E-02	
G4									9.39E-01	
G5-									4.36E-01	
G6-									3.50E-01	
G7-									2.14E-01	
G17–									1.00	

Reaction	pKa Local	pKa Global	Exp		Gas Phase Basicity	Literature Reference
G1+→G3	4.84	3.90	3.2-3.3	G1+→G3	223.31	221.90
G1+→G4	3.65	3.91		G1+→G4	222.66	221.60
G2+→G4	3.61	3.90		G2+→G4	218.31	216.18
G3→G5-	12.26	13.11	9.2-9.6			
G3→G7-	12.57	13.11				
G4→G7-	13.75	13.11				
G4→G6-	13.54	13.11				
G5-→G17–	18.53	18.90				
G7-→G17–	18.23	18.90				
G6-→G17–	18.44	18.90				

Table 6.5: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with Test Set1 Parameters, UFF Radii, α=0.89 for all Ionic Species

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated on geometries optimized at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 1.0 for all species.

6. APPENDIX

Tautomer	$E^0_{(g)}(R)$[1]	ZPE[2] + $\Delta G_{0\rightarrow 298}$	$\Delta G^0_{(g)}(R)$	$E^0_{(g)}(R')$[1]	$\Delta E^0_{(dis)}$	$\Delta G^{0\,R'}_{(sol)}$[3]	$G^0_{(aq)}$	Rel $G^0_{(g)}$	Boltzmann	Rel $G^0_{(aq)}$
	(hartrees)	(hartrees)	(Kcal/mol)	(hartrees)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	Pop.(g)	(Kcal/mol)
G1+	-543.1430	0.0967	-340766.70	-543.1421	0.59	-64.27	-340830.37	0.00	9.96E-01	0.00
G2+	-543.1353	0.0959	-340762.35	-543.1338	0.92	-68.35	-340829.78	4.35	0.64E-03	0.59
G23+	-543.1336	0.0961	-340761.17	-543.1334	0.15	-51.62	-340812.64	5.52	0.89E-04	17.73
G19+	-543.1380	0.0972	-340763.24	-543.1376	0.25	-51.74	-340814.73	3.46	0.29E-02	15.64
G3	-542.7640	0.0836	-340537.11	-542.7624	1.03	-18.3	-340554.38	0.65	2.40E-01	1.39
G4	-542.7652	0.0837	-340537.75	-542.7640	0.71	-18.72	-340555.77	0.00	7.19E-01	0.00
G7a	-542.7625	0.0838	-340536.05	-542.7620	0.32	-7.84	-340543.57	1.70	4.08E-02	10.81
G5-	-542.2105	0.0701	-340198.25	-542.2084	1.31	-70.49	-340267.43	2.82	0.64E-02	-0.10
G6-	-542.2154	0.0705	-340201.08	-542.2138	1.02	-67.28	-340267.34	0.00	7.50E-01	0.00
G7-	-542.2141	0.0705	-340200.25	-542.2114	1.66	-68.48	-340267.06	0.83	1.85E-01	0.27
G34-	-542.2041	0.0695	-340194.61	-542.2031	0.60	-65.81	-340259.82	6.47	0.14E-04	7.52
G15a-	-542.2106	0.0709	-340197.78	-542.2093	0.78	-61.09	-340258.09	3.30	0.29E-02	9.25
G15b-	-542.2136	0.0712	-340199.53	-542.2123	0.83	-58.88	-340257.58	1.54	5.58E-02	9.76
G16-	-542.2036	0.0700	-340194.01	-542.1998	2.41	-72.64	-340264.24	7.07	0.49E-05	3.10
									Pop.(aq)	
G1+									7.30E-01	
G2+									2.70E-01	
G23+									0.11E-07	
G19+									0.36E-06	
G3									8.74E-02	
G4									9.13E-01	
G7a									0.15E-05	
G5-									1.09E-01	
G6-									5.34E-01	
G7-									3.56E-01	
G34-									0.15E-05	
G15a-									0.38E-06	
G15b-									0.16E-06	
G16-									0.16E-02	
Reaction	pKa Local	pKa Global	Exp			Gas Phase Basicity	Literature Reference			
G1+→G3	4.20	3.28	3.2-3.3		G1+→G3	223.31	221.90			
G1+→G4	3.18	3.28			G1+→G4	222.66	221.60			
G2+→G4	2.75	3.28			G2+→G4	218.31	216.18			
G3→G5-	12.23	12.33	9.2-9.6		G23+→G3	217.79	216.40			
G3→G7-	12.50	13.11			G23+→G7a	218.84	217.00			
G4→G7-	13.52	13.11			G19+→G7a	220.91	219.70			
G4→G6-	13.32	13.08								

Table 6.6: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with UFF Radii and α=0.915

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated on geometries optimized at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 1.0 for all species.

Tautomer	$E^0_{(g)}(R)$[1]	ZPE[2]+ $\Delta G_{0\rightarrow 298}$	$\Delta G^0_{(g)}(R)$	$E^0_{(g)}(R')$[1]	$\Delta E^0_{(dis)}$	$\Delta G^{R'}_{(sol)}$[3]	$G^0_{(aq)}$	Rel $G^0_{(g)}$	Boltzmann	Rel $G^0_{(aq)}$
	(hartrees)	(hartrees)	(Kcal/mol)	(hartrees)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	Pop.(g)	(Kcal/mol)
G1+	-543.1430	0.0967	-340766.70	-543.1421	0.59	-60.72	-340826.82	0.00	9.96E-01	0.00
G2+	-543.1353	0.0959	-340762.35	-543.1338	0.92	-64.77	-340826.20	4.35	0.64E-03	0.62
G3	-542.7640	0.0836	-340537.11	-542.7624	1.03	-16.16	-340552.24	0.65	2.40E-01	1.38
G4	-542.7652	0.0837	-340537.75	-542.7640	0.71	-16.57	-340553.62	0.00	7.19E-01	0.00
G5-	-542.2105	0.0701	-340198.25	-542.2084	1.31	-68.41	-340265.35	2.82	0.64E-02	0.00
G6-	-542.2154	0.0705	-340201.08	-542.2138	1.02	-65.23	-340265.29	0.00	7.50E-01	0.07
G7-	-542.2141	0.0705	-340200.25	-542.2114	1.66	-66.36	-340264.94	0.83	1.85E-01	0.41
G17–	-541.5191	0.0575	-339772.26	-541.5171	1.22	-198.14	-339969.18	0.00	1.00	0.00
									Pop.(aq)	
G1+									7.40E-01	
G2+									2.60E-01	
G3									8.87E-02	
G4									9.11E-01	
G5-									4.19E-01	
G6-									3.72E-01	
G7-									2.10E-01	
G17–									1.00	

Reaction	pKa Local	pKa Global	Exp		Gas Phase Basicity	Literature Reference
G1+→G3	3.17	2.25	3.2-3.3	G1+→G3	223.31	221.90
G1+→G4	2.16	2.25		G1+→G4	222.66	221.60
G2+→G4	1.70	2.25		G2+→G4	218.31	216.18
G3→G5-	12.18	12.86	9.2-9.6			
G3→G7-	12.48	12.86				
G4→G7-	13.49	12.86				
G4→G6-	13.24	12.85				
G5-→G17–	18.99	19.37				
G7-→G17–	18.69	19.37				
G6-→G17–	18.94	19.37				

Table 6.7: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with UFF Radii and α=0.915

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated on geometries optimized at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 1.0 for all species.

6. APPENDIX

Tautomer	$E^0_{(g)}(R)$[1]	ZPE[2] + $\Delta G_{0\rightarrow 298}$	$\Delta G^0_{(g)}(R)$	$E^0_{(g)}(R')$[1]	$\Delta E^0_{(dis)}$	$\Delta G^{0\,R'\,3}_{(sol)}$	$G^0_{(aq)}$	Rel $G^0_{(g)}$	Boltzmann	Rel $G^0_{(aq)}$
	(hartrees)	(hartrees)	(Kcal/mol)	(hartrees)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	Pop.(g)	(Kcal/mol)
G1+	-543.1430	0.0967	-340766.70	-543.1421	0.59	-57.99	-340824.09	0.00	9.96E-01	0.00
G2+	-543.1353	0.0959	-340762.35	-543.1338	0.92	-61.94	-340823.37	4.35	0.64E-03	0.72
G3	-542.7640	0.0836	-340537.11	-542.7624	1.03	-14.45	-340550.53	0.65	2.40E-01	1.27
G4	-542.7652	0.0837	-340537.75	-542.7640	0.71	-14.75	-340551.80	0.00	7.19E-01	0.00
G5-	-542.2105	0.0701	-340198.25	-542.2084	1.31	-66.48	-340263.42	2.82	0.64E-02	0.00
G6-	-542.2154	0.0705	-340201.08	-542.2138	1.02	-63.36	-340263.42	0.00	7.50E-01	0.01
G7-	-542.2141	0.0705	-340200.25	-542.2114	1.66	-64.54	-340263.12	0.83	1.85E-01	0.30
G17–	-541.5191	0.0575	-339772.26	-541.5171	1.22	-195.7	-339966.74	0.00	1.00	0.00
									Pop.(aq)	
G1+									7.71E-01	
G2+									2.29E-01	
G3									1.05E-01	
G4									8.95E-01	
G5-									3.87E-01	
G6-									3.80E-01	
G7-									2.33E-01	
G17–									1.00	

Reaction	pKa Local	pKa Global	Exp		Gas Phase Basicity	Literature Reference
G1+→G3	2.42	1.55	3.2-3.3	G1+→G3	223.31	221.90
G1+→G4	1.49	1.56		G1+→G4	222.66	221.60
G2+→G4	0.96	1.55		G2+→G4	218.31	216.18
G3→G5-	12.35	12.91	9.2-9.6			
G3→G7-	12.57	12.91				
G4→G7-	13.49	12.91				
G4→G6-	13.28	12.91				
G5-→G17–	19.36	19.78				
G7-→G17–	19.14	19.78				
G6-→G17–	19.36	19.78				

Table 6.8: Selection of Cavity Model - Raw Data for Deprotonation of Guanine with Test Set4 Parameters, UFF Radii, α=0.98 for all Ionic Species

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated on geometries optimized at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 1.0 for all species.

Tautomer	$E^0_{(g)}(R)$[1]	ZPE[2]+ $\Delta G_{0\rightarrow 298}$	$\Delta G^0_{(g)}(R)$	$E^0_{(g)}(R')$[1]	$\Delta E^0_{(dis)}$	$\Delta G^{0\,R'}_{(sol)}$[3]	$G^0_{(aq)}$	Rel $G^0_{(g)}$	Boltzmann	Rel $G^0_{(aq)}$
	(hartrees)	(hartrees)	(Kcal/mol)	(hartrees)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	Pop.(g)	(Kcal/mol)
G1+	-543.1430	0.0967	-340766.70	-543.1421	0.59	-64.82	-340830.92	0.00	1.00	0.00
G2+	-543.1353	0.0959	-340762.35	-543.1338	0.92	-68.93	-340830.36	4.35	0.64E-03	0.56
G3	-542.7640	0.0836	-340537.11	-542.7624	1.03	-18.51	-340554.59	0.65	0.25	1.52
G4	-542.7652	0.0837	-340537.75	-542.7640	0.71	-19.06	-340556.11	0.00	0.75	0.00
G5-	-542.2105	0.0701	-340198.25	-542.2084	1.31	-75.44	-340272.38	2.82	0.01	0.00
G6-	-542.2154	0.0705	-340201.08	-542.2138	1.02	-72.02	-340272.08	0.00	0.80	0.31
G7-	-542.2141	0.0704	-340200.25	-542.2114	1.66	-73.13	-340271.71	0.83	0.20	0.67
									Pop.(aq)	
G1+									0.72	
G2+									0.28	
G3									0.07	
G4									0.93	
G5-									0.52	
G6-									0.31	
G7-									0.17	
Reaction	pKa Local	pKa Global	Exp			Gas Phase Basicity	Literature Reference			
G1+→G3	4.45	3.45	3.2-3.3		223.31	221.90				
G1+→G4	3.34	3.45			222.66	221.60				
G2+→G4	2.93	3.45			218.31	216.18				
G3→G5-	8.76	9.62	9.2-9.6							
G3→G7-	9.25	9.62								
G4→G7-	10.36	9.62								
G4→G6-	10.09	9.62								

Table 6.9: Calculation of the Local and Global pK_a for Guanine

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

6. APPENDIX

Tautomer	$E^0_{(g)}(R)$[1]	ZPE[2]+ $\Delta G_{0\rightarrow 298}$	$\Delta G^0_{(g)}(R)$	$E^0_{(g)}(R')$[1]	$\Delta E^0_{(dis)}$	$\Delta G^{0\,R'\,3}_{(sol)}$	$G^0_{(aq)}$	Rel $G^0_{(g)}$	Boltzmann	Rel $G^0_{(aq)}$
	(hartrees)	(hartrees)	(Kcal/mol)	(hartrees)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	Pop.(g)	(Kcal/mol)
A1+	-467.8671	0.09395	-293532.08	-467.8667	0.24	-60.43	-293592.27	1.74	0.04	1.55
A2+	-467.8561	0.093036	-293525.77	-467.8557	0.27	-65.56	-293591.06	8.04	0.85E-06	2.76
A3+	-467.8691	0.093977	-293533.34	-467.8688	0.21	-60.69	-293593.82	0.48	0.30	0.00
A4+	-467.8697	0.093788	-293533.82	-467.8694	0.17	-59.85	-293593.49	0.00	0.67	0.33
A5+	-467.8522	0.092461	-293523.70	-467.8517	0.34	-69.49	-293592.85	10.11	0.26E-07	0.97
A6	-467.4959	0.079134	-293308.48	-467.4955	0.29	-9.92	-293318.11	0.00	1.00	0.00
A7	-467.4834	0.080105	-293300.02	-467.4825	0.59	-16.68	-293316.11	8.46	0.6E-06	2.00
A8	-467.4834	0.080152	-293300.00	-467.4829	0.34	-13.98	-293313.64	8.48	0.6E-06	4.48
									Pop.(aq)	
A1+									0.04	
A2+									0.01	
A3+									0.54	
A4+									0.31	
A5+									0.11	
A6									0.97	
A7									0.03	
A8									0.00	
Reaction	pKa Local	pKa Global	Exp							
A1+→A6	2.85	4.24	4.15							
A2+→A6	1.97	4.24								
A2+→A7	3.44	4.24								
A3+→A6	3.99	4.24								
A4+→A7	5.22	4.25								
A5+→A7	4.75	4.24								
A1+→A8	6.14	4.24								

Table 6.10: Calculation of the Local and Global pK_a for Adenine

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

Tautomer	$E^0_{(g)}(R)$[1]	ZPE[2]+ $\Delta G_{0\rightarrow 298}$	$\Delta G^0_{(g)}(R)$	$E^0_{(g)}(R')$[1]	$\Delta E^0_{(dis)}$	$\Delta G^{0\,R'}_{(sol)}$[3]	$G^0_{(aq)}$	Rel $G^0_{(g)}$	Boltzmann	Rel $G^0_{(aq)}$
	(hartrees)	(hartrees)	(Kcal/mol)	(hartrees)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	(Kcal/mol)	Pop.(g)	(Kcal/mol)
C1+	-395.4682	0.0819	-248108.67	-395.4670	0.74	-65.82	-248173.75	0.11	4.55E-01	0.00
C2b+	-395.4686	0.0821	-248108.77	-395.4686	0.03	-56.09	-248164.83	0.00	5.45E-01	8.91
C2+	-395.4562	0.0812	-248101.58	-395.4561	0.11	-61.75	-248163.22	7.20	2.88E-06	10.53
C3b+	-395.4390	0.0800	-248091.52	-395.4386	0.21	-67.47	-248158.78	17.25	1.23E-13	14.97
C3+	-395.4546	0.0814	-248100.41	-395.4544	0.11	-59.87	-248160.17	8.36	4.06E-07	13.58
C2	-395.0913	0.0674	-247881.23	-395.0893	1.24	-17.83	-247897.82	0.00	8.71E-01	0.00
C3	-395.0798	0.0668	-247874.42	-395.0774	1.50	-20.97	-247893.89	6.81	8.90E-06	3.93
C4b	-395.0885	0.0685	-247878.79	-395.0877	0.51	-12.07	-247890.36	2.44	1.42E-02	7.46
C4	-395.0856	0.0683	-247877.14	-395.0849	0.48	-12.84	-247889.51	4.09	8.76E-04	8.31
C5	-395.0899	0.0679	-247880.03	-395.0894	0.32	-8.17	-247887.88	1.21	1.14E-01	9.94
C6b	-395.0577	0.0673	-247860.26	-395.0571	0.39	-13.45	-247873.32	20.97	3.66E-16	24.50
C6	-395.0695	0.0686	-247866.87	-395.0690	0.32	-8.16	-247874.71	14.36	2.57E-11	23.11
C7b	-395.0566	0.0676	-247859.34	-395.0557	0.57	-15.6	-247874.37	21.89	7.77E-17	23.45
C7	-395.0412	0.0660	-247850.75	-395.0401	0.75	-23.37	-247873.37	30.49	3.90E-23	24.45
									Pop.(aq)	
C1+									1.00E+00	
C2b+									2.92E-07	
C2+									1.91E-08	
C3b+									1.07E-11	
C3+									1.11E-10	
C2									9.99E-01	
C3									1.32E-03	
C4b									3.37E-06	
C4									8.05E-07	
C5									5.17E-08	
C6b									1.09E-18	
C6									1.14E-17	
C7b									6.47E-18	
C7									1.19E-18	

Reaction	pKa Local	pKa Global	Exp	Reaction	pKa Local	pKa Global	Exp
C1+→C2	4.15	4.15	4.45	C3+→C3	-2.92	4.16	4.45
C1+→C3	7.03	4.15		C3+→C5	1.49	4.15	
C1+→C4	10.25	4.15		C3+→C6	11.14	4.15	
C1+→C4b	9.62	4.15		C3b+→C3	-3.94	4.16	
C2+→C2	-3.56	4.16		C3b+→C6b	11.14	4.15	
C2+→C5	3.72	4.15		C3b+→C5	0.47	4.16	
C2+→C7	14.36	4.15					
C2+→C2	-2.38	4.16					
C2+→C7b	14.81	4.15					

Table 6.11: Calculation of the Local and Global pK_a for Cytosine

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

6. APPENDIX

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2]+ $\Delta G_{0\rightarrow 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	$\Delta G^{0\,R'}_{(sol)}$[3] (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)	Rel $G^0_{(g)}$ (Kcal/mol)	Boltzmann Pop.(g)	Rel $G^0_{(aq)}$ (Kcal/mol)
T1	-454.3178	0.0822	-285037.17	-454.3168	0.61	-11.10	-285047.67	0.00	1.00E+00	0.00
T6	-454.2732	0.0794	-285010.94	-454.2714	1.13	-22.58	-285032.39	26.24	5.85E-20	15.28
T9	-454.2884	0.0815	-285019.18	-454.2868	1.01	-14.91	-285033.08	18.00	6.40E-14	14.59
T4	-454.3002	0.0820	-285026.23	-454.2995	0.47	-7.64	-285033.40	10.94	9.56E-09	14.27
T7	-454.2884	0.0807	-285019.65	-454.2876	0.50	-12.98	-285032.14	17.52	1.43E-13	15.53
T12	-454.2978	0.0820	-285024.73	-454.2963	0.93	-13.05	-285036.85	12.44	7.56E-10	10.81
T15	-454.2871	0.0814	-285018.37	-454.2854	1.04	-18.68	-285036.01	18.80	1.64E-14	11.66
T10	-454.2775	0.0797	-285013.42	-454.2757	1.15	-17.85	-285030.12	23.75	3.86E-18	17.54
T13	-454.2817	0.0806	-285015.54	-454.2799	1.11	-16.15	-285030.58	21.63	1.38E-16	17.08
T2-	-453.7706	0.0684	-284702.45	-453.7698	0.53	-60.28	-284762.21	0.00	1.00E+00	0.64
T3-	-453.7522	0.0672	-284691.66	-453.7504	1.13	-72.32	-284762.85	10.79	1.22E-08	0.00
T5-	-453.7485	0.0685	-284688.54	-453.7471	0.87	-63.17	-284750.84	13.91	6.35E-11	12.00
T8-	-453.7503	0.0686	-284689.62	-453.7490	0.82	-61.92	-284750.72	12.83	3.92E-10	12.13
T11-	-453.7399	0.0675	-284683.73	-453.7375	1.46	-67.42	-284749.68	18.73	1.87E-14	13.16
T14-	-453.7294	0.0668	-284677.65	-453.7270	1.54	-73.20	-284749.31	24.80	6.58E-19	13.54
									Pop.(aq)	
T1									1.00E+00	
T6									6.33E-12	
T9									2.03E-11	
T4									3.48E-11	
T7									4.12E-12	
T12									1.18E-08	
T15									2.83E-09	
T10									1.38E-13	
T13									2.99E-13	
T2-									2.52E-01	
T3-									7.48E-01	
T5-									1.19E-09	
T8-									9.52E-10	
T11-									1.67E-10	
T14-									8.91E-11	

Reaction	pKa Local	pKa Global	Exp
T1→T2-	11.14	10.54	9.9
T1→T3-	10.67	10.54	
T6→T5-	8.27	10.54	
T6→T3-	-0.53	10.55	
T9→T8-	8.87	10.54	
T9→T3-	-0.02	10.55	
T4→T5-	9.01	10.54	
T4→T2-	0.69	10.55	
T12→T11-	12.39	10.54	
T12→T3-	2.75	10.55	
T15→T14-	12.05	10.54	
T15→T3-	2.12	10.55	
T10→T11-	7.46	10.54	
T10→T2-	-1.72	10.55	
T13→T14-	8.07	10.54	
T13→T2-	-1.38	10.55	
T7→T2-	-0.24	10.55	
T7→T8-	8.18	10.54	

Table 6.12: Calculation of the Local and Global pK_a for Thymine

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2]+ $\Delta G_{0\rightarrow 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	$\Delta G^{0\,R'\,3}_{(sol)}$ (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)	Rel $G^0_{(g)}$ (Kcal/mol)	Boltzmann Pop.(g)	Rel $G^0_{(aq)}$ (Kcal/mol)
8-oxoG1+	-618.3931	0.0973	-387986.53	-618.3899	2.02	-82.64	-388067.15	6.96	7.69E-06	0.00
8-oxoG7+	-618.4007	0.0972	-387991.37	-618.3989	1.14	-70.94	-388061.17	2.12	2.69E-02	5.98
8-oxoG8+	-618.3976	0.0983	-387988.72	-618.3965	0.68	-68.29	-388056.33	4.77	3.09E-04	10.82
8-oxoG9+	-618.4053	0.0984	-387993.49	-618.4034	1.20	-69.28	-388061.57	0.00	9.73E-01	5.58
8-oxoG2	-618.0483	0.0856	-387777.45	-618.0466	1.07	-21.01	-387797.39	0.00	7.80E-01	0.00
8-oxoG5	-618.0474	0.0859	-387776.70	-618.0464	0.59	-13.16	-387789.27	0.75	2.20E-01	8.12
8-oxoG3-	-617.5063	0.0725	-387445.55	-617.5047	1.02	-69.63	-387514.15	0.00	1.00E+00	2.04
8-oxoG4-	-617.4934	0.0705	-387438.73	-617.4909	1.60	-79.06	-387516.19	6.82	1.00E-05	0.00
8-oxoG6-	-617.4971	0.0721	-387440.11	-617.4965	0.42	-67.07	-387506.76	5.44	1.03E-04	9.43
									Pop.(aq)	
8-oxoG1+									1.00E+00	
8-oxoG7+									4.11E-05	
8-oxoG8+									1.17E-08	
8-oxoG9+									8.07E-05	
8-oxoG2									1.00E+00	
8-oxoG5									1.11E-06	
8-oxoG3-									3.09E-02	
8-oxoG4-									9.69E-01	
8-oxoG6-									1.18E-07	
Reaction	pKa Local	pKa Global	Calc Literat.	Exp		Gas Phase Basicity	Literature Reference			
8-oxoG1+→2	-0.37	-0.37	0.22	0.1		202.80	200.80			
8-oxoG8+→2	-8.30	-0.37				204.99	204.20			
8-oxoG7+→2	-4.75	-0.37				207.64	208.50			
8-oxoG7+→5	1.20	-0.37				208.39	208.8			
8-oxoG9+→5	1.49	-0.37				210.51	201.9			
8-oxoG2→3-	9.51	8.00	8.69	8.5						
8-oxoG2→4-	8.02	8.00								

Table 6.13: Calculation of the Local and Global pK_a for 8-oxoguanine

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

7

Appendix

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2]+ $\Delta G_{0\rightarrow 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	$\Delta G^{0\,R'\,3}_{(sol)}$ (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)
Neutrals							
Gh1 (Gh-N3H)	-581.1173	0.0999	-364593.92	-581.1157	0.99	-19.30	-364612.23
Gh2 (Gh-N3imine)	-581.1297	0.1021	-364600.37	-581.1280	1.10	-15.24	-364614.51
Gh3 (Gh-neut-N3Henol-R)	-581.0829	0.0993	-364572.71	-581.0812	1.07	-19.82	-364591.47
Gh4 (Gh-N7C5enol-N3H)	-581.0894	0.1002	-364576.19	-581.0877	1.02	-20.78	-364595.95
Gh5 (Gh-N7C5enol-N3imine)	-581.0955	0.1013	-364579.41	-581.0930	1.60	-20.42	-364598.23
Gh6 (Gh-neut-enol-taut2-R)	-581.0934	0.1002	-364578.75	-581.0916	1.17	-18.89	-364596.47
Gh7 (Gh-N7C8enol-N3H)	-581.0892	0.0990	-364576.88	-	-	-	-
Gh8 (Gh-N7C8enol- N3imine)	-581.1019	0.1022	-364582.82	-581.0998	1.31	-16.65	-364598.16
Gh9 (Gh-zwitter-N7an-N1H2)	-581.0865	0.1002	-364574.47	-581.0792	4.59	-43.90	-364613.78
Cations							
Gh10+ (Gh-N3H-N1H2cat)	-581.5113	0.1153	-364831.55	-581.5083	1.88	-68.41	-364898.08
Gh11+ (Gh-N3H-C5OHcat)	-581.4317	0.1118	-364783.78	-581.4288	1.86	-79.93	-364861.85
Gh12+ (Gh-N3H-C8OHcat)	-581.4414	0.1113	-364790.19	-581.4386	1.79	-77.44	-364865.84
Gh13+ (Gh-enol-N1H2-cat)	-581.4714	0.1120	-364808.57	-581.4684	1.91	-71.21	-364877.87
Gh14+ (Gh-N3imine-C5OHcat)	-581.4567	0.1117	-364799.46	-581.4540	1.65	-64.76	-364862.57
Gh15+ (Gh-N3imine-C8OHcat)	-581.4690	0.1127	-364806.58	-581.4674	0.97	-64.76	-364870.37

Table 7.1: Raw Data Used in the Calculation of the Local and Global pKa for Guanidinohydantoin

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

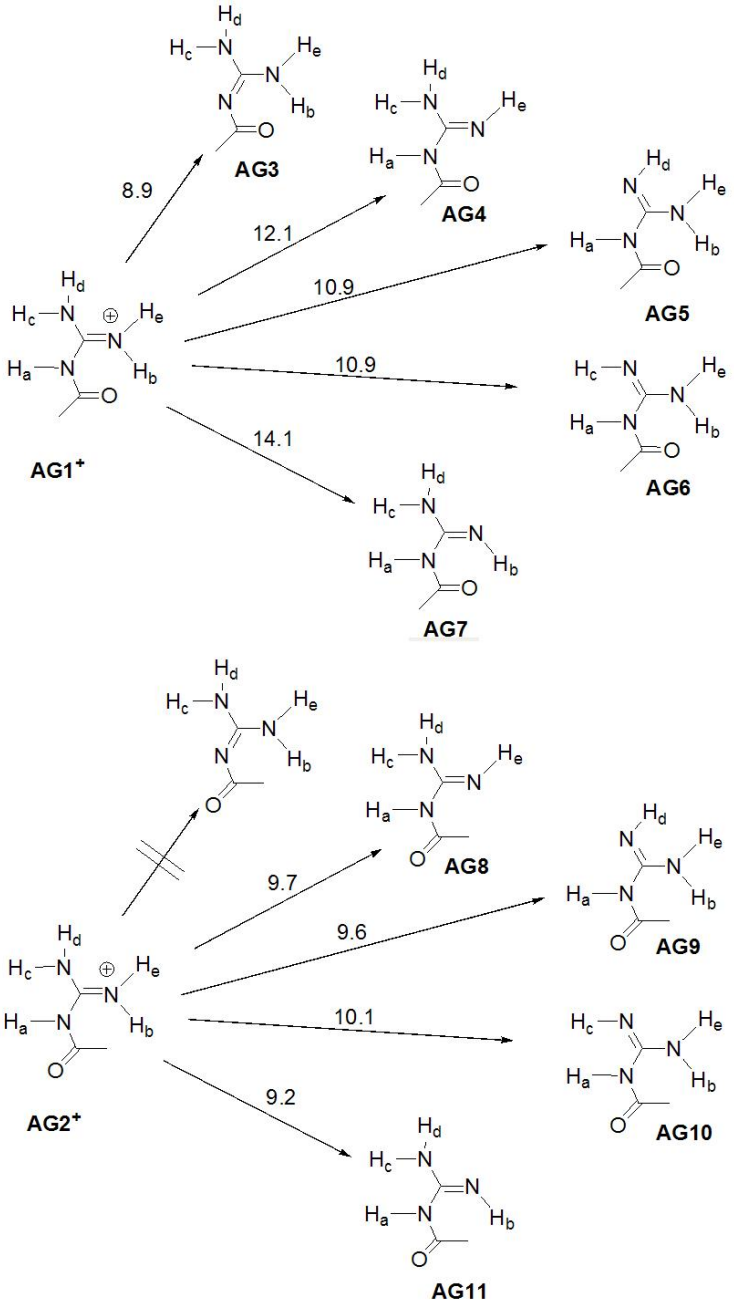


Figure 7.1: Appendix - N-acetylguanidine tautomers and their local pKa values

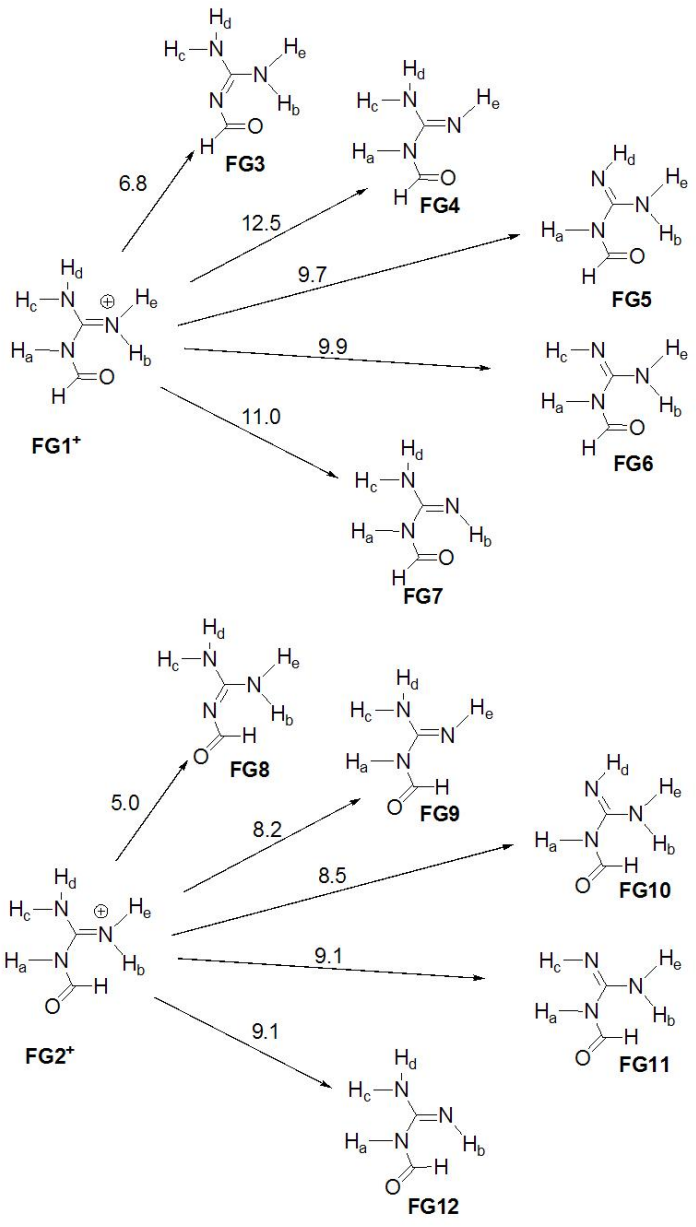


Figure 7.2: Appendix - N-formylguanidine tautomers and their local pKa values

7. APPENDIX

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2]+ $\Delta G_{0\rightarrow 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	$\Delta G^{0\,R'\,3}_{(sol)}$ (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)
Anions							
Gh16- (Gh-N3H-N7-)	-580.5762	0.0881	-364261.83	-580.5712	3.15	-73.60	-364332.28
Gh17- (Gh-N3H-N9-)	-580.5644	0.0865	-364255.39	-580.5616	1.80	-73.00	-364326.60
Gh18- (Gh-N3H-)	-580.5575	0.0865	-364251.01	-580.5552	1.41	-69.48	-364319.08
Gh19- (Gh-N3H-C2NH-)	-580.5327	0.0859	-364235.86	-580.5283	2.75	-82.33	-364315.44
Gh20- (Gh-N3imine-N7b-)	-580.5734	0.0888	-364259.62	-580.5697	2.32	-75.87	-364333.17
Gh21- (Gh-N3imine-N9b-)	-580.5586	0.0861	-364252.02	-580.5539	2.91	-75.65	-364324.76
Gh22- (Gh-enol -)	-580.5439	0.0865	-364242.56	-580.5353	5.45	-80.06	-364317.17
Gh23- (Gh-enol-taut-anion)	-580.5576	0.0874	-364251.15	-580.5554	1.39	-70.96	-364320.15
Gh24- (Gh-N3Henol-N7anion-R)	-580.5209	0.0856	-364228.70	-580.5159	3.15	-82.80	-364308.35
Gh25- (Gh-N3Henol-N9anion-R)	-580.5250	0.0854	-364231.37	-580.5208	2.64	-82.94	-364311.67
Gh26- (Gh-N3Henol-N3anion-R)	-580.5600	0.0872	-364252.18	-580.5566	2.11	-70.27	-364320.34
Gh27- (Gh-N3Henol-C2NHanion-R)	-580.5134	0.0852	-364224.25	-580.5114	1.30	-71.23	-364294.18
Gh28- (Gh-enol-taut2-N7anion)	-580.5325	0.0874	-364234.85	-580.5267	3.64	-78.63	-364309.84
Gh29- (Gh-enol-taut2-N9anion)	-580.5296	0.0866	-364233.49	-580.5265	1.96	-76.62	-364308.15
Gh 30a-(Gh-enol-taut2-C2NHa-anion)	-580.5298	0.0868	-364233.47	-580.5265	2.02	-67.95	-364299.40
Gh 30b- (Gh-enol-taut2-C2NHb-anion)	-580.5168	0.0860	-364225.79	-580.5120	2.99	-76.11	-364298.91
Gh 30c- (Gh-enol-taut2-C2NHc-anion)	-580.5341	0.0866	-364236.36	-580.5325	1.02	-64.86	-364300.20
Gh 30d- (Gh-enol-taut2-C2NHd-anion)	-580.5250	0.0869	-364230.41	-580.5212	2.39	-70.93	-364298.96

Table 7.2: Raw Data Used in the Calculation of the Local and Global pKa for Guanidinohydantoin

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

Reaction		pKa Local	pKa Global
Gh10+→Gh1	Gh+,N3H	11.42	9.63
Gh10+→Gh2	Gh+,N3imine	9.75	9.64
Gh10+→Gh3	Gh+,Gh enol tauto	22.40	9.63
Gh13+→Gh3	Gh-cat-N3Henol-R,Gh-neut-N3Henol-R	11.83	9.63
Gh13+→Gh6	Gh-cat-N3Henol-R,Gh-neut-enol-taut2-R	8.16	9.63
Gh11+→Gh4	Gh-C5OHcat-N3H,Gh-N7C5enol-N3H	-3.20	9.63
Gh15+→Gh8	Gh-C8OHcat-N3imine,Gh-N7C8enol-N3imine	1.43	9.64
Gh14+→Gh5	Gh-C5OHcat-N3imine,Gh-N7C5enol-N3imine	-4.34	9.64
Gh10+→Gh9	Gh+,Ghzwitter	10.29	9.63
Gh1→Gh19-	N3HC2NH-,N3H	19.45	8.14
Gh1→Gh16-	N3HN7-,N3H	7.10	8.15
Gh1→Gh17-	N3HN9-,N3H	11.27	8.14
Gh1→Gh18-	N3H-,N3H	16.78	8.15
Gh2→Gh18-	N3H-,N3imine	18.45	8.15
Gh2→Gh20-	N3imineN7b-,N3imine	8.12	8.15
Gh2→Gh21-	N3imineN9b-,N3imine	14.29	8.15
Gh6→Gh26-	GhenolN3-,Ghenoltaut	4.88	8.15
Gh6→Gh29-	GhenoltautN9-,Ghenoltaut	13.78	8.15
Gh6→Gh23-	Ghenoltaut-,Ghenoltaut	5.02	8.15
Gh3→Gh22-	Gh-enol-anion,Gh-neut-N3Henol-R	2.96	8.15
Gh3→Gh27-	Gh-N3Henol-C2NHanion-R,Gh-neut-N3Henol-R	19.81	8.14
Gh3→Gh24-	Gh-N3Henol-N7anion-R,Gh-neut-N3Henol-R	9.42	8.14
Gh3→Gh25-	Gh-N3Henol-N9anion-R,Gh-neut-N3Henol-R	6.99	8.15
Gh3→Gh26-	Gh-N3Henol-N3anion-R,Gh-neut-N3Henol-R	0.63	8.15
Gh6b→Gh26-	Gh-N3Henol-N3anion-R,Gh-neut-enol-taut2-R	4.30	8.15
Gh6b→Gh29-	Gh-enol-taut2-N9anion,Gh-neut-enol-taut2-R	13.23	8.15
Gh6b→Gh24-	Gh-enol-taut2-N7anion,Gh-neut-enol-taut2-R	12.00	8.15
Gh6b→Gh27a-	Gh-enol-taut2-C2NHa-anion,Gh-neut-enol-taut2-R	19.65	8.15
Gh6b→Gh27b-	Gh-enol-taut2-C2NHb-anion,Gh-neut-enol-taut2-R	20.01	8.15
Gh6b→Gh27c-	Gh-enol-taut2-C2NHc-anion,Gh-neut-enol-taut2-R	19.06	8.15
Gh6b→Gh27d-	Gh-enol-taut2-C2NHd-anion,Gh-neut-enol-taut2-R	19.97	8.15
Gh4→Gh16-	Gh-N7C5enol-N3H,Gh-N3H-N7an	-4.83	8.16
Gh8→Gh20-	Gh-N7C8enol-N3imine,Gh-N3imine-N7b-anion	-3.86	8.15
Gh5→Gh20-	Gh-N7C5enol-N3imine,Gh-N3imine-N7b-anion	-3.82	8.15
Gh9→Gh16-	GhZwitter,Gh N3H-N7an	8.24	8.15

Table 7.3: Local pKa for Deprotonation of Various Guanidinohydantoin Tautomers

7. APPENDIX

Reaction	pKa Local	pKa Global	Exp	Reaction	pKa Local	pKa Global	Exp
Guan+→Guan	14.53	14.53	13.7				
FG1+→FG3	6.82	6.72	——	AG1+→AG3b	8.88	8.54	8.23
FG1+→FG4	12.55	6.72		AG1+→AG4	12.06	8.54	
FG1+→FG5	9.70	6.72		AG1+→AG5	10.86	8.54	
FG1+→FG6	9.91	6.72		AG1+→AG6	10.88	8.54	
FG1+→FG7	10.99	6.72		AG1+→AG7	14.09	8.54	
FG2+→FG8	4.97	6.72		AG2+→AG3a	3.31	8.54	
FG2+→FG9	8.18	6.72		AG2+→AG8	9.65	8.54	
FG2+→FG10	8.49	6.72		AG2+→AG9	9.62	8.54	
FG2+→FG11	9.05	6.72		AG2+→AG10	10.10	8.54	
FG2+→FG12	9.12	6.72		AG2+→AG11	9.19	8.54	

Table 7.4: Calculation of the Local and Global pKa for Guanidine, Formylguanidine, and Acetylguanidine

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2]+ $\Delta G_{0\to 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	$\Delta G^{0\,R'}_{(sol)}$[3] (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)	Rel $G^0_{(g)}$ (Kcal/mol)	Boltzmann Pop.(g)	Rel $G^0_{(aq)}$ (Kcal/mol)
Hydantoin	-376.8663	0.0498	-236455.95	-376.8653	0.66	-12.56	-236467.86	0.00	1.00E+00	0.00
Hydantoin N3-	-376.3094	0.0381	-236113.82	-376.3076	1.12	-69.93	-236182.63	0.00	1.00E+00	0.00
Hydantoin N1-	-376.3010	0.0371	-236109.19	-376.3000	0.59	-67.92	-236176.51	4.63	4.04E-04	6.12
Phenytoin	-839.1199	0.1982	-526431.34	-839.1190	0.56	-1.63	-526432.41	0	1.00E+00	0.00
Phenytoin N3-	-838.5724	0.1849	-526096.13	-838.5698	1.67	-55.77	-526150.22	0.00	9.99E-01	0.00
Phenytoin N1-	-838.5655	0.1844	-526092.11	-838.5619	2.29	-54.06	-526143.89	4.02	1.14E-03	6.34
Hydantoin									Pop.(aq)	
Hydantoin N3-									1.00E+00	
Hydantoin N1-									1.00E+00	
Phenytoin									3.28E-05	
Phenytoin N3-									1.00E+00	
Phenytoin N3-									1.00E+00	
Phenytoin N1-									2.26E-05	
Reaction	pKa Local	pKa Global	Exp							
H→HN3-	10.97	10.97	9.1							
H→HN1-	15.45	10.97								
Ph→PhN3-	8.72	8.72	8.31							
Ph→PhN1-	13.36	8.72								

Table 7.5: Calculation of the Local and Global pKa for Hydantoin and Phenyltoin

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2]+ $\Delta G_{0\rightarrow 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	$\Delta G^{0\,R'\,3}_{(sol)}$ (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)
Neutrals							
Sp1 (Sp-neut-N3H)	-693.2972	0.0880	-434995.33	-693.2933	2.40	-26.18	-435019.10
Sp2 (Sp-neut-N1H)	-693.2940	0.0871	-434993.95	-693.2921	1.20	-22.63	-435015.38
Sp3 (Sp-neut-C2imina-b)	-693.2939	0.0880	-434993.28	-693.2927	0.77	-19.56	-435012.07
Sp4 (Sp-neut-C2imina-a)	-693.2953	0.0883	-434993.98	-693.2936	1.03	-19.37	-435012.32
Sp5 (Sp-neut-N1H-N7C5-R2)	-693.2654	0.0873	-434975.82	-693.2628	1.60	-24.98	-434999.20
Sp6 (Sp-neut-N1H-N7C8-R2)	-693.2650	0.0872	-434975.64	-693.2621	1.80	-25.56	-434999.40
Sp7 (Sp-neut-N1H-N9C8-R2)	-693.2648	0.0871	-434975.6	-693.2630	1.14	-22.59	-434997.05
Sp8 (Sp-neut-N1H-N1C6-R)	-693.2777	0.0880	-434983.09	-693.2765	0.72	-17.68	-435000.05
Sp9 (Sp-neut-N3H-N7C5-R2)	-693.2674	0.0879	-434975.42	-693.2635	1.20	-28.68	-435002.90
Sp10 (Sp-neut-N3H-N7C8-R)	-693.2699	0.0881	-434975.05	-693.2643	0.40	-28.53	-435003.17
Sp11 (Sp-neut-N3H-N9C8-R)	-693.2678	0.0877	-434977.13	-693.2644	2.18	-26.62	-435001.57
Sp12 (Sp-neut-C2imine-N7C5-R)	-693.2648	0.0875	-434975.34	-693.2629	1.18	-21.85	-434996.01
Sp13 (Sp-neut-C2imine-N7C8-R)	-693.2663	0.0883	-434975.8	-693.2642	1.33	-21.77	-434996.24
Sp14 (Sp-neut-C2imine-N9C8-R)	-693.2658	0.0874	-434976.04	-693.2645	0.80	-19.67	-434994.92
Sp15 (Sp-neut-C2imine-N1C6-R)	-693.2715	0.0871	-434979.76	-693.2701	0.86	-18.10	-434997.00
Cations							
Sp16+ (Spiro+)	-693.6567	0.1005	-435213.15	-693.6531	2.27	-79.19	-435290.07
Sp17+ (Spiro C5OH+)	-693.6216	0.0989	-435192.09				
Sp18+ (Spiro C8OH+)	-693.6299	0.0987	-435197.43	-693.6283	1.04	-71.35	-435267.74
Sp19+ (Spiro C6OH+)	-693.6128	0.0990	-435186.51	-693.6103	1.61	-79.13	-435264.03

Table 7.6: Raw Data Used in the Calculation of the Local and Global pKa for Spiroiminodihydantoin

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

7. APPENDIX

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2] + $\Delta G_{0\rightarrow 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	Dip.Momen. in sol.	$\Delta G^{0\,R'\,3}_{(sol)}$ (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)
Anions								
Sp20- (Sp-N7-anion)	-692.7564	0.0756	-434663.78	-692.7492	4.51	19.3	-83.07	-434742.33
Sp21- (Sp-N9a-anion)	-692.7445	0.0744	-434657.09	-692.7396	3.06	15.2	-83.04	-434737.07
Sp22- (Sp-N3H-anion)	-692.7627	0.0751	-434668.07	-692.7606	1.31	6.2	-70.94	-434737.69
Sp23a- (Sp-C210a-anion)	-692.7579	0.0752	-434664.98	-692.7548	1.94	12.4	-72.05	-434735.09
Sp23b- (Sp-C210b-anion)	-692.7646	0.0756	-434668.93	-692.7624	1.43	8.5	-67.44	-434734.94
Sp24- (Sp-N1H-N7b-anion)	-692.7442	0.0743	-434656.93	-692.7387	3.46	20.9	-84.33	-434737.80
Sp25- (Sp-N1H-N9b-anion)	-692.7338	0.0738	-434650.7	-692.7296	2.63	15.2	-83.80	-434731.87
Sp26- (Sp-C2N10-N7c-anion-a)	-692.7617	0.0754	-434667.24	-692.7586	1.99	12.4	-71.17	-434736.42
Sp27- (Sp-C2N10-N9c-a-anion)	-692.7528	0.0747	-434662.07	-692.7509	1.18	7.3	-70.07	-434730.96
Sp28- (Sp-C2N10-N3d-anion)	-692.7443	0.0738	-434657.32	-692.7431	0.75	5.6	-71.24	-434727.81
Sp29- (Sp-C2N10-N7c-anion-b)	-692.7589	0.0752	-434665.63	-692.7561	1.76	13.7	-72.46	-434736.33
Sp30- (Sp-C2N10-N9c-b-anion)	-692.7498	0.0743	-434660.45	-692.7482	1.02	8.1	-71.64	-434731.08
Sp31- (Sp-C2N10-N3c-anion)	-692.7534	0.0746	-434662.51	-692.7519	0.90	4.6	-66.33	-434727.94

Table 7.7: Raw Data Used in the Calculation of the Local and Global pKa for Spiroiminodihydantoin

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

Reaction		pKa Local	pKa Global
Sp16+→Sp4	C2imA,Cat	5.49	0.52
Sp16+→Sp3	C2imB,Cat	5.67	0.52
Sp16+→Sp2	N1H,Cat	3.25	0.52
Sp16+→Sp1	N3H,Cat	0.52	0.52
Sp1→Sp20-	N7-anion,neut-N3H	4.77	4.77
Sp1→Sp21-	N9a-anion,neut-N3H	8.62	4.77
Sp1→Sp23a-	C210a,neut-N3H	10.08	4.77
Sp1→Sp23b-	C210b,neut-N3H	10.19	4.76
Sp1→Sp22-	N3H-anion,neut-N3H	8.17	4.77
Sp2→Sp22-	N3H-anion,neut-N1H	5.44	4.77
Sp2→Sp24-	N1H-N7b-anion,neut-N1H	5.37	4.77
Sp2→Sp25-	N1H-N9b-anion,neut-N1H	9.71	4.77
Sp2→Sp28-	C2N10-N3c-anion,neut-N1H	12.59	4.77
Sp4→Sp28-	C2N10-N3c-anion,neut-C2imine-a	10.35	4.77
Sp4→Sp29-	C2N10-N7c-anion-a,neut-C2imine-a	4.13	4.77
Sp3→Sp26-	C2N10-N7c-anion-b,neut-C2imine-b	4.02	4.77
Sp4→Sp28-	C2N10-N9c-anion-a,neut-C2imine-a	8.13	4.77
Sp3→Sp27-	C2N10-N9c-anion-b,neut-C2imine-b	7.87	4.77
Sp4→Sp23a-	C210a-anion-a,neut-C2imine-a	5.11	4.77
Sp3→Sp23b-	C210a-anion-b,neut-C2imine-b	5.04	4.76
Sp5→Sp24-	N1H-N7b-anion,neut-N1H-N7C5-R2	-6.49	4.77
Sp6→Sp24-	N1H-N7b-anion,neut-N1H-N7C8-R2	-6.34	4.77
Sp7→Sp25-	N9a-anion,neut-N1H-N9C8-R2	-7.54	4.77
Sp8→Sp22-	N3H-anion,neut-N1H-N1C6-R	-5.80	4.77
Sp9→Sp20-	N7-anion,neut-N3H-N7C5-R2	-7.11	4.77
Sp10→Sp20-	N7-anion,neut-N3H-N7C8-R	-6.91	4.77
Sp11→Sp21-	N9a-anion,neut-N3H-N9C8-R	-4.22	4.77
Sp12→Sp26-	C2N10-N7c-anion-b,neut-C2imine-N7C5-R-b	-7.75	4.77
Sp14→Sp27-	C2N10-N9c-anion-b,neut-C2imine-N9C8-R	-4.70	4.77
Sp15→Sp28-	C210b-anion,neut-C2imine-N1C6-R	-6.01	4.77
Sp13→Sp29-	C2N10-N7c-anion-b,neut-C2imine-N7C8-R	-7.58	4.77

Table 7.8: Local pKa for Deprotonation of Various Spiroiminodihydantoin Tautomers

7. APPENDIX

Tautomer	$E^0_{(g)}(R)$[1] (hartrees)	ZPE[2]+ $\Delta G_{0\rightarrow 298}$ (hartrees)	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$E^0_{(g)}(R')$[1] (hartrees)	$\Delta E^0_{(dis)}$ (Kcal/mol)	$\Delta G^{0\,R'\,3}_{(sol)}$ (Kcal/mol)	$G^0_{(aq)}$ (Kcal/mol)
Neutrals							
Sp-neut-N1H-CH2	-619.2369	0.1081	-388509.2	-619.2352	1.09	-17.69	-388525.80
Sp-neut-N3H-CH2	-619.2340	0.1068	-388508.21	-619.2324	1.00	-16.88	-388524.09
Sp-neut-C2imine-a-CH2	-619.2294	0.1070	-388505.22	-619.2288	0.38	-18.36	-388523.20
Sp-neut-C2imine-b-CH2	-619.2287	0.1075	-388504.44	-619.2273	0.85	-18.70	-388522.30
Sp-N7an-N1H-CH2	-618.6702	0.0916	-388163.95	-618.6671	1.95	-78.74	-388240.74
Sp-N7an-N3H-CH2	-618.6860	0.0951	-388171.66	-618.6824	2.27	-68.25	-388237.64
Sp-N7an-C2imine-a-CH2	-618.6882	0.0950	-388173.13	-618.6861	1.33	-65.89	-388237.69
Sp-N7an-C2imine-b-CH2	-618.6840	0.0945	-388170.78	-618.6822	1.18	-68.06	-388237.66
Sp-neut-N1H-C6CH2	-657.3378	0.1102	-412416.5	-657.3356	1.33	-17.65	-412432.83
Sp-neut-N3H-C6CH2	-657.3409	0.1111	-412417.98	-657.3393	1.04	-16.05	-412432.99
Sp-neut-C2imine-a-C6CH2	-657.3365	0.1104	-412415.64	-657.3354	0.71	-15.14	-412430.07
Sp-neut-C2imine-b-C6CH2	-657.3354	0.1101	-412415.16	-657.3344	0.66	-15.65	-412430.15
Sp-neut-N1H-C6CH2-N7an	-656.7810	0.0976	-412075.06	-656.7756	3.42	-80.43	-412152.07
Sp-neut-N3H-C6CH2-N7an	-656.7953	0.0986	-412083.44	-656.7911	2.62	-72.89	-412153.71
Sp-C2imine-a-C6CH2-N7an	-656.7972	0.0976	-412085.23	-656.7948	1.50	-67.81	-412151.54
Sp-C2imine-b-C6CH2-N7an	-656.7942	0.0974	-412083.5	-656.7923	1.18	-69.36	-412151.68
Sp-neut-C2imina-a	-693.2953	0.0883	-434993.98	-693.2936	1.03	-19.37	-435012.32
Sp-neut-C2imina-b	-693.2939	0.0880	-434993.28	-693.2927	0.77	-19.56	-435012.07
Sp-neut-N1H	-693.2940	0.0871	-434993.95	-693.2921	1.20	-22.63	-435015.38
Sp-neut-N3H	-693.2972	0.0880	-434995.33	-693.2933	2.40	-26.18	-435019.10
Sp-N1H-N7b-anion	-692.7442	0.0743	-434656.93	-692.7387	3.46	-84.33	-434737.80
Sp-N3H-N7-anion	-692.7564	0.0756	-434663.78	-692.7492	4.51	-83.07	-434742.34
Sp-C2N10-N7c-anion-a	-692.7617	0.0754	-434667.24	-692.7586	1.99	-71.17	-434736.42
Sp-C2N10-N7c-anion-b	-692.7589	0.0752	-434665.63	-692.7561	1.76	-72.46	-434736.33

Table 7.9: Raw Data for Evaluation of Through-Space Electronic Effects on the pKa of Sp

1. Gas phase calculations conducted at B3LYP/aug-cc-pVTZ//B3LYP/6-31+G(d,p) and B3LYP/aug-cc-pVTZ//IEF-PCM/B3LYP/6-31+G(d,p) for the gas and solution phase optimized geometries, respectively.

2. Frequency calculation conducted at B3LYP/6-31+G(d,p) on geometry optimized in the gas phase at the same level of theory.

3. Free energy of solvation calculated at IEF-PCM/B3LYP/6-31+G(d,p) using UFF atom model α= 0.91 for cation and neutral species and 0.83 for anions on the geometry optimized in solution with radii=UFF and α= 1.0 for all species.

Reaction	pKa Local	pKa Global	$\Delta G^0_{(g)}(R)$ (Kcal/mol)	$\Delta G^{0\,R'}_{(sol)}$ (Kcal/mol)	$\Delta E^0_{(dis)}$ (Kcal/mol)
CH2 SpC2imAneut→C2imAN7an	11.18	10.87	325.815557	-47.53	0.94
CH2 SpC2imBneut→C2imBN7an	10.53	10.87	327.380412	-49.36	0.33
CH2 N1Hneut→N1HN7an	10.85	10.87	338.967689	-61.05	0.87
CH2 N3Hneut→N3HN7an	11.87	10.87	330.27017	-51.37	1.28
C6CH2 SpC2imAneut→C2imAN7an	6.06	6.81	324.124657	-52.67	0.80
C6CH2 SpC2imBneut→C2imBN7an	6.02	6.81	325.378856	-53.71	0.52
C6CH2 N1Hneut→N1HN7an	7.69	6.81	335.166894	-62.78	2.09
C6CH2 N3Hneut→N3HN7an	6.61	6.81	328.261209	-56.84	1.58
Sp SpC2imAneut→C2imAN7an	4.13	4.77	320.460112	-51.80	0.96
Sp SpC2imBneut→C2imBN7an	4.02	4.77	321.368854	-52.90	1.00
Sp N1Hneut→N1HN7an	5.37	4.77	330.74101	-61.70	2.26
Sp N3Hneut→N3HN7an	4.77	4.77	325.26364	-56.89	2.11

Table 7.10: Site-Specific pKa for Evaluation of Through-Space Electronic Effects on the pKa of Sp

INTERMEDIATE	Free Enenrgy[1] BS-0	Free Enenrgy[1] BS-I	Free Enenrgy[1] BS-II	Free Enenrgy[1] BS-III	Free Enenrgy[1] BS-IV	ZPE[2]
1	-1269.966876	-1269.915517	-2495.220761	-2495.553584	-2495.596723	0.51282
III	-1254.840888	-1254.789377	-2034.858774	-2035.190052	-2035.225002	0.51427
IV	-1240.445008	-1240.390551	-1575.223685	-1575.554152	-1575.588412	0.52077
i-PrO$^-$	-193.795520	-193.760143	-193.760143	-193.812427	-193.857950	0.09256
Acetone	-193.117830	-193.095342	-193.095342	-193.153662	-193.166186	0.08393
Cl$^-$	-15.113325	-15.118302	-460.365658	-460.365658	-460.378759	0.00000
IX	-1433.605361	-1433.488926	-1768.322486	-1768.700843	-1768.739075	0.61110
Acetophenone	-384.839167	-384.796473	-384.796473	-384.895260	-384.913992	0.13891
PhenilEthanol	-386.044945	-385.996438	-385.996438	-386.098715	-386.120791	0.16203
TS2-Si	-1625.284164	-1625.151736	-1959.985543	-1960.411960	-1960.461060	0.65632
TS2-Re	-1625.279728	-1625.146186	-1959.979234	-1960.406181	-1960.455337	0.65604
i-PrOH	-194.325441	-194.300099	-194.300099	-194.361176	-194.376113	0.10801
I	-1239.230432	-1239.181058	-1574.014929	-1574.339445	-1574.373425	0.49935
TS1	-1433.538696	-1433.458371	-1768.292866	-1768.676900	-1768.720386	0.60264
TS3	-1433.562525	-1433.446245	-1768.280015	-1768.665033	-1768.705383	0.60435
X					-2034.7566	0.49910

Table 7.11: Single point energy computed at five different basis set for each intermediate into hydrogen transfer Ru assisted

1. Energy in Hartree comprehensive of electrostatic, repulsive, dispersive and cavitation energy of the intermediates of Path I and II

2. Zero point energy correction in Hartree computed on the gas phase optimized geometry

TRANSITION STATE	Number of iaginary mode	Frequency cm^{-1}
TS1	1	-559.8
TS2	1	-888.4
TS3-Si	1	-939.7
TS3-Re	1	-960.2
TS4	1	-52.7
TS5	1	-143.9

Table 7.12: Number of imaginary modes and relative frequencies in cm^{-1} of the most important transition states found in Path I,II and III.

INTERMEDIATE	Free Enenrgy[1] BS-IV	ZPE[2]
XI	-1768.267949	0.5958
XII	-1575.1126	0.5050
TS4	-1768.223189	0.5918
TS5	-1959.967274	0.6442

Table 7.13: Most important intermediate and transition state energy of Path III.

1. Energy in Hartree comprehensive of electrostatic, repulsive, dispersive and cavitation energy of the intermediate of Path III

2. Zero point energy correction in Hartree computed on the gas phase optimized geometry

Gas Phase											
Mol.	Freq.[a]	Rh	S1	S2	N1	N2	N3	C	O	P1	P2
c1	2077.19	-0.217029	-0.092648	-0.092647	-0.382179	-0.833	-0.382179	0.9262	0.9262	0.341057	-0.289251
c2	2113.01	-0.191207	-0.039424	0.062448	-0.354517	-0.838058	-0.413416	0.925137	0.89705	0.339608	-0.24947
c3	2144.71	-0.151055	0.101354	0.101354	-0.406034	-0.845922	-0.406034	0.894126	0.894126	0.334836	-0.214684
CH_2Cl_2											
Mol.	Freq.[a]	Rh	S1	S2	N1	N2	N3	C	O	P1	P2
c1	2027.01	-0.228611	-0.164918	-0.164317	-0.363765	-0.857629	-0.363818	0.927002	0.927308	0.347331	-0.327229
c2	2057.78	-0.192687	-0.111763	0.046473	-0.359809	-0.861	-0.408411	0.926442	0.895601	0.343051	-0.294543
c3	2090.92	-0.157682	0.077655	0.078958	-0.406406	-0.86186	-0.4007	0.895749	0.896024	0.338163	-0.264026

Table 7.14: Mulliken populations (a.u.) for selected atoms, in the gas phase and in solution (PCM, CH_2Cl_2) for **c1**, **c2** and **c3**. For compound **c2**, the protonation is on the N1 atom [B3LYP/BS-0//B3LYP/BS-I; BS-0=6-31g** + Lanl2DZ(Rh), BS-I=6-31g** + dgdzvp(Rh)]. [a]CO frequencies stretching are reported in cm^{-1}.

References

[1] JULIA FAHRENKAMP-UPPENBRINK, PHIL SZUROMI, JAKE YESTON, and ROBERT COONTZ. **8**. *SCIENCE*, **321**, 2008. 1

[2] Richard A. Friesner. **66486653**. *PNAS*, **102**, 2005. 1

[3] Hohenberg P. and Kohn W. **B864**. *Phys. Rev.*, **1326**, 1964. 4, 9

[4] Sham L. J. and Kohn W. **561**. *Phys. Rev.*, **145**, 1964. 4

[5] Kohn W., Becke A. D., and R. G. Parr. **12974**. *J. Phys.Chem.*, **100**, 1996. 4

[6] Johnson B. G., Gill P. M. W., and J. A. Pople. **5612**. *J. Chem. Phys.*, **98**, 1993. 4, 5

[7] Becke A. D. **3098**. *Phys. Rev.A*, **38**, 1998. 4

[8] Becke A. D. **5648**. *J. Chem. Phys.*, **98**, 1998. 4, 5

[9] Lee C. T., Yang W. T., and Parr R. G. **785**. *Phys.Rev. B*, **37**, 1988. 4, 5

[10] Bartlett R. J. **1697**. *J. Phys. Chem.*, **93**, 1989. 5

[11] Martin J. M. L. and G. de Oliveira. **1843**. *J. Chem.Phys.*, **111**, 1988. 5

[12] Thomas J. R., Deeleeuw B. J. Vacek, G. Crawford T. D., Yamaguchi Y., and Schaefer H. F. **403**. *J. Chem. Phys.*, **99**, 1993. 5

[13] Grimme S. **9095**. *J. Chem. Phys.*, **118**, 2003. 5

[14] Tsuzuki S., Uchimaru T., and Tanabe K. **202**. *Chem. Phys. Lett.*, **287**, 1998. 5, 7

[15] Tsuzuki S., Uchimaru T., Matsumura K., Mikami M., and Tanabe K. **11906**. *J. Chem. Phys.*, **110**, 1999. 5

[16] Murphy R. B., Beach M. D., Friesner R. A., and Ringnalda M. N. **1481**. *J. Chem. Phys.*, **103**, 1995. 5

[17] Roos B. O., Andersson K., Fulscher M. P., Malmqvist P. A., Serrano-Andres L., Pierloot K., and Merchan M. **219**. *Adv. Chem. Phys.*, **93**, 1996. 6

[18] Dutta A. and Sherrill C. D. **1610**. *J. Chem. Phys.*, **118**, 2003. 6

[19] Murray C. W., Laming G. J., Handy N. C., and Amos R. D. **551**. *Chem. Phys. Lett.*, **199**, 1992. 6

[20] Redfern P. C., Zapol P., Curtiss L. A., Rajh T., and Thurnauer M. C. **11419**. *J. Phys. Chem. B*, **107**, 2003. 6

[21] Curtiss L. A., Raghavachari K., Redfern P. C., and Pople J. A. **1063**. *J. Chem. Phys.*, **106**, 1997. 6

[22] Guallar V., Jacobson M., McDermott A., and Friesner R. A. **227**. *J. Mol. Biol.*, **337**, 2004. 6

[23] Stratmann R. E., Scuseria G. E., and Frisch M. J. **109**. *J. Chem. Phys.*, **8218**, 1998. 7

[24] Field M. J., Bash P. A., and Karplus M. **700**. *J. Comput. Chem.*, **11**, 1990. 7

[25] Tannor D. J., Marten B., Murphy R., Friesner R. A., Sitkoff D., Nicholls A., Ringnalda M., Goddard W. A., and Honig B. **11875**. *J. Am. Chem. Soc.*, **116**, 1994. 8

[26] Tomasi J. and Persico M. **2027**. *Chem. Rev.*, **94**, 1994. 8

REFERENCES

[27] Honig B. and Nicholls A. **1144**. *Science*, **268**, 1995. 8

[28] Klamt A., Jonas V., Burger T., and Lohrenz J. C. W. **5074**. *J. Phys. Chem. A*, **102**, 1998. 8

[29] Barone V., Cossi M., and Tomasi J. **404**. *J. Comput. Chem.*, **19**, 1998. 8

[30] Cortis C. M., Langlois J. M., Beachy M. D., and Friesner R. A. **5472**. *J. Chem. Phys.*, **105**, 1996. 8

[31] Marten B., Kim K., Cortis C., Friesner R. A., Murphy R. B., Ringnalda M. N., Sitkoff D., and Honig B. **11775**. *J. Phys. Chem.*, **100**, 1996. 8

[32] Klicic J. J., Friesner R. A., Liu S. Y., and Guida W. C. **1327**. *J. Phys. Chem. A*, **106**, 2002. 8

[33] Kohn W. and Sham L. J. **140A**. *Phys. Rew.*, **1133**, 1965. 10

[34] Kohn W. and Sham L. J. **137A**. *Phys. Rew.*, **1697**, 1965. 10

[35] Slater J. C. **81**. *Phys. Rev.*, **385**, 1951. 11

[36] Ceperly D. M. and Alden B. J. **45**. *Phys. Rev. Let*, **566**, 1980. 11

[37] Becke A. D. **23**. *J. Quantum Chem.*, **1915**, 1983. 11

[38] Becke A. D. **84**. *J. Chem. Phys.*, **4524**, 1986. 11

[39] Becke A. D. **38**. *Phys. Rev. A*, **3098**, 1988. 11

[40] Becke A. D. **98**. *J. Chem. Phys.*, **1372**, 1993. 12

[41] Becke A. D. **98**. *J. Chem. Phys.*, **143**, 1993. 12

[42] Ulig H. H. **1215**. *J. Phys. Chem.*, **41**, 1937. 15

[43] Hummer G., Garde S., Garcia A. E., Paulaitis M. E., and Pratt L. R. **8951**. *Proc. Natl. Acad. Sci. USA*, **93**, 1996. 15

[44] Pierotti R. A. **712**. *Chem. Rev.*, **76**, 1976. 15

[45] Vigne'-Maeder F. and Claverie P. **24**. *J. Am. Chem. Soc.*, **109**, 1987. 15

[46] Amovilli C. and Mennucci B. **1051**. *J. Phys. Chem. B*, **101**, 1997. 17

[47] Lee B. and Richards F. M. **The interpretation of protein structures: Estimation of static accessibility**. *J. Mol. Biol.*, **55**, 1971. 19

[48] Richards F. M. **151**. *Ann. Rev. Biophys. Eng.*, **6**, 1977. 19

[49] J. L. Pascual-Ahuir. **Un metodo para calculo de superficies moleculares**. *GEPOL Thesis Doctoral*, **Univerist de Valencia**, 1988. 20

[50] Benedetta Mennucci and Roberto Cammi. *Continuum Solvation Models in Chemical Physics*. Jon Wiley and Sons, West Sussex PO19 8SQ England, 1st edition edition, 2008. 21

[51] V. Schettino, R. Bini, M. Ceppatelli, L. Ciabini, and M. Citroni. **131**. *Adv.Chem. Phys.*, **105**, 2005. 23

[52] H.K. Mao, D.R. Herschbach, and in: R. Manaa (Ed.). **189**. *Chemistry Under Extreme Conditions*, 2005. 23

[53] Y. Song, Z. Liu, R.J. Hemsley, M. Ho-Kwang, and D.R. Herschbah. **122**. *J. Chem. Phys.*, **174511**, 2005. 23

[54] M. Mugnai, G. Cardini, and V. Schettino. **70**. *Phys. Rev. B*, 2004. 23, 30, 54

[55] P.F. McMillan and Nat. Mater. **1**. *Nat. Mater.*, **18**, 2002. 23

[56] M. Citroni, M. Ceppatelli, R. Bini, and V. Schettino. **295**. *Science*, **2002**, 2002. 23

[57] R.J. Hemsley. **51**. *Ann. Rev. Phys. Chem.*, **763**, 2000. 23

[58] M. Ceppatelli, M. Santoro, R. Bini, and V. Schettino. **113**. *J. Chem. Phys.*, **5991**, 2000. 23

[59] D. Ben-Amotz, M.R. Zakin, H.E. King, and D.R. Herschbach. **92**. *J. Phys. Chem.*, **1392**, 1988. 23

[60] J. Tomasi, B. Mennucci, and Chem R. Cammi. **105**. *Chem. Rev.*, **2999**, 2005. 23, 25, 30, 35

[61] E.P. Wigner and H. Huntington. **3**. *J. Chem. Phys.*, **764**, 1935. 24

[62] A. Michels, J. De Boers, and A. Bijl. **4**. *Physica*, **981**, 1937. 24

[63] A. Sommerfeld and H. Welker. **32**. *Ann. Phys.*, **56**, 1938. 24

[64] S.R. de Groot and C.A. ten Seldam. **12**. *Physica*, **669**, 1946. 24

[65] R.W. Zwanzig. **18**. *J. Chem. Phys.*, **1412**, 1950. 24

[66] C.A. ten Seldam and S.R. de Groot. **18**. *Physica*, **891,905,910**, 1952. 24

[67] W. Byers Brown. **28**. *J. Chem. Phys.*, **522**, 1958. 24

[68] J. de Boer. **15**. *Physica*, **843**, 1949. 24

[69] R.J. Riddell and G.E.W. Uhlenbeck. **18**. *J. Chem. Phys.*, **1068**, 1950. 24

[70] G. Marc and W.G. McMillan. **58**. *Adv. Chem. Phys.*, **209**, 1985. 24

[71] E.V. Ludena. **69**. *J. Chem. Phys.*, **1770**, 1978. 24

[72] E.V. Ludena and M. Gregori. **71**. *J. Chem. Phys.*, **2235**, 1979. 24

[73] R. LeSar and D.R. Herschbach. **85**. *J. Phys. Chem.*, **2798**, 1981. 24, 25

[74] R. LeSar and D.R. Herschbach. **87**. *J. Phys. Chem.*, **5202**, 1983. 24

[75] J. Gorecki and W. Byers Brown. **89**. *J. Chem. Phys.*, **2138**, 1988. 24

[76] J.P. Connerade and R. Smaoune. **33**. *J. Phys. B*, **3467**, 2000. 24

[77] K.D. Sen, B. Mayer, P.C. Schmidt, J. Garza, R. Vragas, and A. Vela. **90**. *Int. J. Quantum Chem.*, **491**, 2002. 24

[78] S.A. Cruz and J. Soullard. **391**. *Chem. Phys. Lett.*, **138**, 2004. 24

[79] R. Cammi and J. Tomasi. **101**. *J. Chem. Phys.*, **3888**, 1994. 25

[80] B. Mennucci and R. Cammi Eds. **3**. *Continuum Solvation Models in Chemical Physics*, 2008. 25

[81] J. Tomasi, B. Mennucci, R. Cammi, and in: S. Wilson (Ed.). **3**. *Handbook of Molecular Physical and Quantum Chemistry*, 2001. 26

[82] R. Cammi and J. Tomasi. **100**. *J. Chem. Phys.*, **7495**, 1994. 26

[83] R. Cammi and J. Toma. **101**. *J. Chem. Phys.*, **3888**, 1994. 26

[84] E. Cances and B. Mennucci. **109**. *J. Chem. Phys.*, **249**, 1998. 26

[85] E. Cances and B. Mennucci. **109**. *J. Chem. Phys.*, **260**, 1998. 26

[86] B. Mennucci, R. Cammi, and J. Tomasi. **110**. *J. Chem. Phys.*, **6858**, 1999. 26

[87] M. Cossi, G. Scalmani, N. Rega, and B. Barone. **117**. *J. Chem. Phys.*, **43**, 2002. 26

[88] F.M. Floris, J. Tomasi, and J.L. Pascual-Ahuir. **12**. *J. Comput. Chem.*, **784**, 1991. 26

[89] M. Cossi, J. Tomasi, and R. Cammi. **592**. *Int. J. Quantum Chem.*, **695**, 1995. 26

[90] C. Curutchet, M. Orozco, F.J. Luque, B. Mennucci, and J. Tomasi. **27**. *Comput. Chem.*, **1769**, 1006. 27

[91] C. Amovilli and B. Mennucci. **101**. *J. Phys. Chem. B*, **1051**, 1997. 27, 28

[92] R. Cammi, V. Verdolino, B. Mennucci, and J. Tomasi. **135**. *Chemical Physics*, **344**, 2008. 29

REFERENCES

[93] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. **Gaussian, Inc., Wallingford, CT**. *Gaussian 03 revision C.02*, 2004. 30, 31, 36, 38

[94] A.D. Becke. **98**. *J. Chem. Phys.*, **5648**, 1993. 30, 31, 35, 38

[95] W.J. Hehre, R. Ditchfield, and J.A. Pople. **56**. *J. Chem. Phys.*, **2257**, 1972. 30

[96] A. Bondi. **68**. *J. Phys. Chem.*, **441**, 1964. 30

[97] E. Cances, B. Mennucci, and J. Tomasi. **107**. *J. Chem. Phys.*, **3032**, 1997. 30, 45

[98] D. Chipman. **106**. *J. Chem. Phys.*, **10194**, 1997. 30

[99] E. Cances and B. Mennucci. **114**. *J. Chem. Phys.*, **4744**, 2001. 30

[100] J.-F. Ma F.-F. Li and, J. Yang, H.-Q. Jia, and N.-H. Hu. **787**. *J. Mol. Struct.*, **106**, 2006. 31

[101] Y.Q. Wei, K.C. Wu, B.T. Zhuang, and Z.F. Zhou. **751**. *J. Mol. Struct.*, **133**, 2005. 31

[102] M. B. Ferrari, F. Bisceglie, E. Cavalli, G.o Pelosi, P. Tarasconi, and V. Verdolino. **360**. *Inorganica Chimica Acta*, **3233**, 2007. 31

[103] E. Drent, P. Arnoldy, and P. H. M. Budzelaar. **455**. *J. Organomet. Chem.*, **247**, 1993. 33

[104] P. Denis, A. Jean, J. F. Croizy, A. Mortreux, and F. Petit. **112**. *J. Am. Chem. Soc.*, **1292**, 1990. 33

[105] T. B. Rauchfuss, J. L. Clements, S. F. Angew, and D. M. Roundhill. **16**. *Inorg. Chem.*, **775**, 1977. 33

[106] M. Koprovski, R.-S. Sebastin, V. Maraval, M. Zablocka, V. Cadierno, B. Donnadieu, A. Igau, A.-M. Caminade, and J.-P. Majoral. **21**. *Organometallics*, **4680**, 2002. 33

[107] J.-D. Huang, X.-P. Hu, S.-B. Yu, J. Deng, D.-Y. Wang, and Z.-C. Duanand Z. Zheng. **270**. *J. Mol. Catal. A Chem.*, **127**, 2007. 33

[108] P. Cheruku, S. Gohil, and P. G. Andersson. **9**. *Org. Lett.*, **1659**, 2007. 33

[109] J. Andrieu, J. M. Camus, P. Richard, R. Poli, L. Gonsalvi, F. Vizza, and M. Peruzzini. **2006**. *Eur. J. Inorg. Chem.*, 51. 33

[110] C. J. Adams, M. I. Bruce, B. W. Skelton, and A. H. White. **513**. *J. Organomet. Chem.*, **255**, 1996. 33

[111] L. Crociani, R. Anacardio, P. Traldi, and B. Corain. **282**. *Inorg. Chim. Acta*, **119**, 1998. 33

[112] L. Crociani, F. Tisato, F. Refosco, G. Bandoli, and B. Corain. **1689**. *Eur. J. Inorg. Chem.*, 1998. 33

[113] M. B. Smith and A. M. Z. Slawin. **24**. *New J. Chem.*, **65**, 2000. 33

[114] O. Crespo, E. J. Fernndez, M. Gil, M. C. Gimeno, P. G. Jones, A. Laguna, J. M. Lpez de Luzuriaga, and M. E. Olmos. **1319**. *J. Chem. Soc. Dalton Trans.*, 2002. 33

[115] L.-C. Liang. **250**. *Coord. Chem. Rev.*, **1152**, 2006. 33

[116] K. R. Reddy, K. Surekha, G.-H. Lee, S. M. Peng, and S.-T. Liu. **19**. *Organometallics*, **2637**, 2000. 33

[117] L. Dahlenburg, K. Herbst, and A. Zahl. **616**. *J. Organomet. Chem.*, **19**, 2000. 33

[118] K. R. Reddy, W. W. Tsai, K. Surekha, G.-H. Lee, S.-M. Peng, J.-T. Chen, and S.-T. Liu. **1776**. *J. Chem. Soc., Dalton Trans.*, 2002. 33

[119] H.-Y. Wang and G.-X. Jin. **1665**. *Eur. J. Inorg. Chem.*, 2005. 33

[120] R. Morris, A. Habtemariam, Z. Guo, S. Parsons, and P. J. Sadler. **339**. *Inorg. Chim. Acta*, **551**, 2002. 33

[121] M. A. Rankin, R. McDonald, M. J. Ferguson, and M. Stradiotto. **24**. *Organometallics*, **4981**, 2005. 33

[122] C. G. Arena, S. Calamia, C. Graiff F. Faraone, and A. Tiripicchio. **3149**. *J. Chem. Soc., Dalton Trans.*, 2000. 33

[123] K. Mauthner, C. Slugovc, K. Mereiter, R. Schmid, and K. Kirchner. **16**. *Organometallics*, **195**, 1997. 33

[124] C. Standfest-Hauser, C. Slugovc, K. Mereiter, R. Schmid, K. Kirchner, L. Xiao, and J. Weissensteiner. **2989**. *J. Chem. Soc., Dalton Trans.*, 2001. 34

[125] M. Ito, A. Osaku, S. Kitahara, and T. Ikariya M. Hirakawa. **44**. *Tetrahedron Lett.*, **7521**, 2003. 34

[126] C. Thoumazet, M. Melaimi, L. Ricard, F. Mathey, and P. Le Floch. **22**. *Organometallics*, **1580**, 2003. 34

[127] M. Ito, M. Hirakawa, A. Osaku, and T. Ikariya. **22**. *Organometallics*, **4190**, 2003. 34

[128] M. Ito, S. Kitahara, and T. Ikariya. **127**. *J. Am. Chem. Soc.*, **6172**, 2005. 34

[129] M. Ito, A. Sakaguchi, C. Kobayashi, and T. Ikariya. **129**. *J. Am. Chem. Soc.*, **290**, 2007. 34

[130] M. Ito, A. Osaku, A. Shiibashi, and T. Ikariya. **9**. *Org. Lett.*, **1821**, 2007. 34

[131] R. J. Lundgren, M. A. Rankin, R. McDonald, G. Schatte, and M. Stradiotto. **46**. *Angew. Chem. Int. Ed.*, **4732**, 2007. 34

[132] G. Zassinovich, G. Mestroni, and S. Gladiali. **92**. *Chem. Rev.*, **1051**, 1992. 34

[133] D. A. Alonso, P. Brandt, S. J. M. Nordin, and P. G. Andersson. **121**. *J. Am. Chem. Soc.*, **9580**, 1999. 34, 35

[134] M. Yamakawa and R. Noyori M. Ito. **122**. *J. Am. Chem. Soc.*, **1466**, 2000. 34, 64

[135] R. L. Chowdhury and J. E. Bckvall. **1063**. *J. Chem. Soc., Chem. Commun.*, 1991. 34

[136] A. Aranyos, G. Csjernyik, and J.E. Bckvall K. J. Szab. **351**. *Chem. Commun.*, 1999. 34

[137] O. Pmies and J. E. Bckvall. **7**. *Chem. Eur. J.*, **5052**, 2001. 34

[138] J. R. Fulton, A. W. Holland, D. J. Fox, and R. G. Bergman. **35**. *Acc. Chem. Res.*, **44**, 2002. 34

[139] H. E. Bryndza and W. Tam. **88**. *Chem. Rev.*, **1163**, 1988. 34

[140] W. Baratta, G. Chelucci, S. Gladiali, K. Siega, M. Toniutti, M. Zanette, E. Zangrando, and P. Rigo. **44**. *Angew. Chem. Int. Ed.*, **6214**, 2005. 34

[141] W. Baratta, K. Siega, P. Rigo, M. Zanette, E. Zangrando, and P. Rigo. **13**. *Chem. Eur. J.*, **7479**, 2007. 34

[142] M. Yamakawa, I. Yamada, and R. Noyori. **40**. *Angew. Chem. Int. Ed.*, **2818**, 2001. 35, 64, 69

[143] R. Noyori, M. Yamakawa, and S. Hashiguchi. **66**. *J. Org. Chem.*, **7931**, 2001. 35, 64, 69

[144] P. Brandt, P. Roth, and P. G. Andersson. **69**. *J. Org. Chem.*, **4885**, 2004. 35, 69

[145] A. Comas-Vives, G. Ujaque, and A. Lleds. **26**. *Organometallics*, **4135**, 2007. 35, 69

[146] D. G. I. Petra, J. N. H. Reek, J. W. Handgraaf, E. J. Meijer, P. Dierkes, P. C. Kamer, J. Brussee, H. E. Schoemaker, and P. W. N. M. van Leeuwen. **6**. *Chem. Eur. J.*, **2818**, 2000. 35, 108

[147] W. J. Hehre, L. Radom, P. van R. Schleyer, and J. A. Pople. **Wiley-Interscience**. *Ab Initio Molecular Orbital Theory*, 2005. 35

[148] Feliz M., Freixa Z., van Leeuwent P. W. N. M., and C. Bo. **24**. *Organometallics*, **5718**, 2005. 37

[149] Doppiu A., Englert U., and Salzer A. **2166**. *Chem. Commun.*, 2004. 37

[150] Kinnunen T. and Laasonen K. **665**. *J. Organomet. Chem.*, **150**, 2003. 37

[151] Martin H. C., James N. H., Aitken J., Gaunt J. A., Adams H., and Haynes A. **22**. *Organometallics*, **4451**, 2003. 37

[152] Gonsalvi L., Adams H., Sunley G. J., Ditzel E., and Haynes A. **124**. *J. Am. Chem. Soc.*, **13597**, 2002. 37

[153] Delferro M., Cauzzi D., Pattacini R., Tegoni M., Graiff C., and Tiripicchio A. **2302**. *Eur. J. Inorg. Chem.*, 2008. 37, 72

[154] Godbout N., Salahub D. R., Andzelm J., and Wimmer E. **70**. *Can. J. Chem.*, **560**, 1992. 38

[155] Sosa C., Andzelm J., Elkin B. C., Wimmer E., Dobbs K. D., and Dixon D. A. **96**. *J. Phys. Chem.*, **6630**, 1992. 38

[156] Reed A. E., Curtiss L. A., and Weinhold F. **88**. *Chem. Rev.*, **899**, 1988. 38

[157] Gimisis T. and Cismas C. **Isolation, characterization, and independent synthesis of guanine oxidation products**. *European Journal of Organic Chemistry*, **6**:1351–1378, 2006. 39

REFERENCES

[158] Burrows C. J. and Muller J. G. **Oxidative nucleobase modifications leading to strand scission**. *Chemical Reviews*, **98**(3):1109–1151, 1998. 39

[159] Wachters A. J. H. **Gaussian Basis Sets for Molecular Wavefunctions Containing Third Row Atoms**. *J. Chem. Phys.*, **52**:1033, 1970. 39

[160] Pratviel G. and Meunier B. **Guanine oxidation: One- and two-electron reactions**. *Chemistry-a European Journal*, **12**(23):6018–6030, 2006. 39

[161] Sayre L. M., Perry G., and Smith M. A. **Oxidative stress and neurotoxicity**. *Chemical Research in Toxicology*, **21**:172–188, 2008. Times Cited: 0. 39

[162] Beckman K. B. and Ames B. N. **Oxidative decay of DNA**. *Journal of Biological Chemistry*, **272**(32):19633–19636, 1997. 39

[163] Foksinski M., Rozalski R., Guz J., Ruszkowska B., Sztukowska P., Piwowarski M., Klungland A., and Olinski R. **Urinary excretion of DNA repair products correlates with metabolic rates as well as with maximum life spans of different mammalian species**. *Free Radical Biology and Medicine*, **37**(9):1449–1454, 2004. 39

[164] Steenken S., Jovanovic S. V., Bietti M., and Bernhard K. **The trap depth (in DNA) of 8-oxo-7,8-dihydro-2 ' deoxyguanosine as derived from electron-transfer equilibria in aqueous solution**. *Journal Of The American Chemical Society*, **122**(10):2373–2374, 2000. 39, 40

[165] Luo W. C., Muller J. G., Rachlin E. M., and Burrows C. J. **Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model**. *Chemical Research In Toxicology*, **14**(7):927–938, 2001. 39

[166] Luo W. C., Muller J. G., Rachlin E. M., and Burrows C. J. **Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-oxo-7,8-dihydroguanosine**. *Organic Letters*, **2**(5):613–616, 2000. 39

[167] Niles J. C., Wishnok J. S., and Tannenbaum S. R. **Spiroiminodihydantoin and guanidinohydantoin are the dominant products of 8-oxoguanosine oxidation at low fluxes of peroxynitrite: Mechanistic studies with O-18**. *Chemical Research In Toxicology*, **17**(11):1510–1519, 2004. 39

[168] Ye Y., Muller J. G., Luo W. C., Mayne C. L., Shallop A. J., Jones R. A., and Burrows C. J. **Formation of C-13-, N-15-, and O-18-labeled guanidinohydantoin from guanosine oxidation with singlet oxygen. Implications for structure and mechanism**. *Journal Of The American Chemical Society*, **125**(46):13926–13927, 2003. 39

[169] Suzuki T., Friesen M. D., and Ohshima H. **Identification of products formed by reaction of 3 ',5 '-di-O-acetyl-2 '-deoxyguanosine with hypochlorous acid or a myeloperoxidase-H2O2-Cl- system**. *Chemical Research in Toxicology*, **16**(3):382–389, 2003. 39

[170] Munk B. H., Burrows C. J., and Schlegel H. B. **An Exploration of Mechanisms for the Transformation of 8-Oxoguanine to Guanidinohydantoin and Spiroiminodihydantoin by Density Functional Theory**. *Journal of the American Chemical Society*, **130**:5245–5256, 2008. 39

[171] Lehninger, editor. *Principles of Biochemistry*. W. H. Freeman and Company, New York, fourth edition, 2005. 39

[172] Chen J. G., McAllister M. A., Lee J. K., and Houk K. N. **Short, strong hydrogen bonds in the gas phase and in solution: Theoretical exploration of pKa matching and environmental effects on the strengths of hydrogen bonds and their potential roles in enzymatic catalysis**. *Journal of Organic Chemistry*, **63**(14):4611–4619, 1998. 39, 40

[173] Acharya P., Cheruku P., Chatterjee S., Acharya S., and J. Chattopadhyaya. **Measurement of nucleobase pK(a) values in model mononucleotides shows RNA-RNA duplexes to be more stable than DNA-DNA duplexes**. *Journal of the American Chemical Society*, **126**(9):2862–2869, 2004. 39, 84

[174] Acharya S., Barman J., Cheruku P., Chatterjee S., Acharya P., Isaksson J., and J. Chattopadhyaya. **Significant pK(a) perturbation of nucleobases is an intrinsic property of the sequence context in DNA and RNA**. *Journal of the American Chemical Society*, **126**(28):8674–8681, 2004. 39

[175] Steenken S. and Jovanovic S. V. **How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution**. *Journal Of The American Chemical Society*, **119**(3):617–618, 1997. 39, 40

[176] Candeias L. P. and Steenken S. **STRUCTURE AND ACID-BASE PROPERTIES OF ONE-ELECTRON-OXIDIZED DEOXYGUANOSINE, GUANOSINE, AND 1-METHYLGUANOSINE**. *Journal of the American Chemical Society*, **111**(3):1094–1099, 1989. Times Cited: 195. 39, 40

[177] Roitzsch M. and Lippert B. **Metal coordination and imine-amine hydrogen bonding as the source of strongly shifted adenine pK(a) values**. *Journal of the American Chemical Society*, **126**(8):2421–2424, 2004. 39

[178] Chatterje S., Pathmasiri W., Plashkevych O., Honcharenko D., Varghese O. P., Maiti M., and Chattopadhyaya J. **The chemical nature of the 2 '-substituent in the pentose-sugar dictates the pseudoaromatic character of the nucleobase (pK(a)) in DNA/RNA**. *Organic and Biomolecular Chemistry*, **4**(9):1675–1686, 2006. Times Cited: 2. 39, 41

[179] Rogstad K. N., Jang Y. H., Sowers L. C., and Goddard W. A. **First principles calculations of the pK(a) values and tautomers of isoguanine and xanthine**. *Chemical Research In Toxicology*, **16**(11):1455–1462, 2003. 40, 41, 78, 100

[180] Jang Y. H., Hwang S. G., and Chung D. S. **DFT calculation of site-specific acid dissociation constants of purine nucleobases**. *Chemistry Letters*, **36**:1496–1497, 2007. Times Cited: 0. 40, 41

[181] Jang Y. H., Goddard W. A., Noyes K. T., Sowers L. C., Hwang S., and Chung D. S. **pK(a) values of guanine in water: Density functional theory calculations combined with Poisson-Boltzmann continuum-solvation model**. *Journal Of Physical Chemistry B*, **107**(1):344–357, 2003. 40, 41, 43, 78, 79, 81, 83, 90, 100

[182] Chandra A. K., Nguyen M. T., and Zeegers-Huyskens T. **Theoretical study of the interaction between thymine and water. Protonation and deprotonation enthalpies and comparison with uracil**. *Journal of Physical Chemistry A*, **102**(29):6010–6016, 1998. Times Cited: 76. 40

[183] Chandra A. K., Michalska D., Wysokinsky R., and Zeegers-Huyskens T. **Theoretical study of the acidity and basicity of the cytosine tautomers and their 1 : 1 complexes with water**. *Journal of Physical Chemistry A*, **108**(44):9593–9600, 2004. Times Cited: 14. 40

[184] Chandra A. K., Nguyen M. T., Uchimaru T., and Zeegers-Huyskens T. **Protonation and deprotonation enthalpies of guanine and adenine and implications for the structure and energy of their complexes with water: Comparison with uracil, thymine, and cytosine**. *Journal of Physical Chemistry A*, **103**(44):8853–8860, 1999. Times Cited: 115. 40

[185] Chandra A. K., Nguyen M. T., and Zeegers-Huyskens T. **Theoretical study of the protonation and deprotonation of cytosine. Implications for the interaction of cytosine with water**. *Journal of Molecular Structure*, **519**:1–11, 2000. Times Cited: 31. 40

[186] Podolyan Y., Gorb L., and Leszczynski J. **Protonation of nucleic acid bases. A comprehensive post-Hartree-Fock study of the energetics and proton affinities**. *Journal of Physical Chemistry A*, **104**(31):7346–7352, 2000. Times Cited: 50. 40

[187] Giese B. and McNaughton D. **Density functional theoretical (DFT) and surface-enhanced Raman spectroscopic study of guanine and its alkylated derivatives - Part 1. DFT calculations on neutral, protonated and deprotonated guanine**. *Physical Chemistry Chemical Physics*, **4**(20):5161–5170, 2002. Times Cited: 18. 40

[188] Tang C. L., Alexov E., Pyle A. M., and Honig B. **Calculation of pK(a)s in RNA: On the structural origins and functional roles of protonated nucleotides**. *Journal of Molecular Biology*, **366**(5):1475–1496, 2007. 40, 41

[189] Jang Y. H., Goddard W. A., Noyes K. T., Sowers L. C., Hwang S., and Chung D. S. **First principles calculations of the tautomers and pK(a) values of 8-oxoguanine: Implications for mutagenicity and repair**. *Chemical Research In Toxicology*, **15**(8):1023–1035, 2002. 40, 41, 43, 82, 83, 84, 90, 100

[190] Kornyushyna O., Berges A. M., Muller J. G., and Burrows C. J. **In vitro nucleotide misinsertion opposite the oxidized guanosine lesions spiroiminodihydantoin and guanidinohydantoin and DNA synthesis past the lesions using Escherichia coli DNA polymerase I (Klenow fragment)**. *Biochemistry*, **41**(51):15304–15314, 2002. 40

[191] Henderson P. T., Delaney J. C., Gu F., Tannenbaum S. R., and Essigmann J. M. **Oxidation of 7,8-dihydro-8-oxoguanine affords lesions that are potent sources of replication errors in vivo**. *Biochemistry*, **41**(3):914–921, 2002. 40

[192] Henderson P. T., Delaney J. C., Muller J. G., Neeley W. L., Tannenbaum S. R., C. J. Burrows, and J. M. Essigmann. **The hydantoin lesions formed from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors in vivo**. *Biochemistry*, **42**(31):9257–9262, 2003. 40

[193] Duarte V., Muller J. G., and Burrows C. J. **Insertion of dGMP and dAMP during in vitro DNA synthesis opposite an oxidized form of 7,8-dihydro-8-oxoguanine**. *Nucleic Acids Research*, **27**(2):496–502, 1999. Times Cited: 106. 40

[194] Chinyengetere F. and Jamieson E. R. **Impact of the oxidized guanine lesion spiroiminodihydantoin on the conformation and thermodynamic stability of a 15-mer DNA duplex**. *Biochemistry*, **47**:2584–2591, 2008. Times Cited: 0. 40

[195] Jia L., Shafirovich V., Shapiro R., Geacintov N. E., and Broyde S. **Spiroiminodihydantoin lesions derived from guanine oxidation: Structures, energetics, and functional implications**. *Biochemistry*, **44**(16):6043–6051, 2005. 40

[196] Lim C., Bashford D., and Karplus M. **Absolute Pka Calculations With Continuum Dielectric Methods**. *Journal Of Physical Chemistry*, **95**(14):5610–5620, 1991. 40, 44

[197] Hwang S., Jang Y. H., and Chung D. S. **Gas Phase Proton Affinity, Basicity, and pKa Values for Nitrogen Containing Heterocyclic Aromatic Compounds**. *Bulletin of the Korean Chemical Society*, **26**(4):585–588, 2005. 40

[198] Tissandier M. D., Cowen K. A., Feng W. Y., Gundlach E., Cohen M. H., A. D. Earhart, J. V. Coe, and T. R. Tuttle. **The proton's absolute aqueous enthalpy and Gibbs free energy of solvation from cluster-ion solvation data**. *Journal of Physical Chemistry A*, **102**(40):7787–7794, 1998. Times Cited: 172. 40

[199] Sadlej-Sosnowska N. **Calculation of acidic dissociation constants in water: solvation free energy terms. Their accuracy and impact**. *Theoretical Chemistry Accounts*, **118**(2):281–293, 2007. Times Cited: 1. 40

[200] Bryantsev V. S., Diallo M. S., and Goddard W. A. **pK(a) calculations of aliphatic amines, diamines, and aminoamides via density functional theory with a Poisson-Boltzmann continuum solvent model**. *Journal of Physical Chemistry A*, **111**(20):4422–4430, 2007. Times Cited: 2. 40

[201] Burk P., Koppel I. A., Koppel I., Leito I., and Travnikova O. **Critical test of performance of B3LYP functional for prediction of gas-phase acidities and basicities**. *Chemical Physics Letters*, **323**(5-6):482–489, 2000. Times Cited: 39. 40

[202] Barone V., Improta R., and Rega N. **Computation of protein pK's values by an integrated density functional theory/Polarizable Continuum Model approach**. *Theoretical Chemistry Accounts*, **111**(2-6):237–245, 2004. Times Cited: 31. 40

[203] Farras P., Teixidor F., and Branchadell V. **Prediction of pK(a) values of nido-carboranes by density functional theory methods**. *Inorganic Chemistry*, **45**(19):7947–7954, 2006. Times Cited: 2. 40

REFERENCES

[204] Ervin K. M. and DeTuro V. F. **Anchoring the gas-phase acidity scale**. *Journal of Physical Chemistry A*, **106**(42):9947–9956, 2002. Times Cited: 53. 40

[205] Kelly C. P., Cramer C. J., and Truhlar D. G. **Adding explicit solvent molecules to continuum solvent calculations for the calculation of aqueous acid dissociation constants**. *Journal of Physical Chemistry A*, **110**(7):2493–2499, 2006. Times Cited: 24. 40

[206] Liptak M. D., Gross K. C., Seybold P. G., Feldgus S., and Shields G. C. **Absolute pK(a) determinations for substituted phenols**. *Journal Of The American Chemical Society*, **124**(22):6421–6427, 2002. 40

[207] Liptak M. D. and Shields G. C. **Experimentation with different thermodynamic cycles used for pK(a) calculations on carboxylic acids using complete basis set and gaussian-n models combined with CPCM continuum solvation methods**. *International Journal Of Quantum Chemistry*, **85**(6):727–741, 2001. Sp. Iss. SI. 40

[208] Liptak M. D. and Shields G. C. **Accurate pK(a) calculations for carboxylic acids using Complete Basis Set and Gaussian-n models combined with CPCM continuum solvation methods**. *Journal Of The American Chemical Society*, **123**(30):7314–7319, 2001. 40

[209] Liptak M. D. and Shields G. C. **Comparison of density functional theory predictions of gas-phase deprotonation data**. *International Journal Of Quantum Chemistry*, **105**(6):580–587, 2005. 40

[210] Merrill G. N. and Kass S. R. **Calculated gas-phase acidities using density functional theory: Is it reliable?** *Journal of Physical Chemistry*, **100**(44):17465–17471, 1996. Times Cited: 70. 40

[211] Pliego J. R. **Thermodynamic cycles and the calculation of pK(a)**. *Chemical Physics Letters*, **367**(1-2):145–149, 2003. 40

[212] Pliego J. R. and Riveros J. M. **Theoretical calculation of pK(a) using the cluster-continuum model**. *Journal Of Physical Chemistry A*, **106**(32):7434–7439, 2002. 40

[213] Swart M. and Bickelhaupt F. M. **Proton affinities of anionic bases: Trends across the periodic table, structural effects, and DFT validation**. *Journal of Chemical Theory and Computation*, **2**(2):281–287, 2006. Times Cited: 6. 40

[214] da Silva C. O., da Silva E. C., and Nascimento M. A. C. **Ab initio calculations of absolute pK(a) values in aqueous solution I. Carboxylic acids**. *Journal of Physical Chemistry A*, **103**(50):11194–11199, 1999. Times Cited: 61. 40, 42

[215] Silva C. O., da Silva E. C., and Nascimento M. A. C. **Ab initio calculations of absolute pK(a) values in aqueous solution II. Aliphatic alcohols, thiols, and halogenated carboxylic acids**. *Journal of Physical Chemistry A*, **104**(11):2402–2409, 2000. Times Cited: 76. 40

[216] Takano Y. and Houk K. N. **Benchmarking the conductor-like polarizable continuum model (CPCM) for aqueous solvation free energies of neutral and ionic organic molecules**. *Journal Of Chemical Theory And Computation*, **1**:70–77, 2005. 40

[217] Saracino G. A. A., Improta R., and Barone V. **Absolute pK(a) determination for carboxylic acids using density functional theory and the polarizable continuum model**. *Chemical Physics Letters*, **373**(3-4):411–415, 2003. 40

[218] Schuurmann G., Cossi M., Barone V., and Tomasi J. **Prediction of the pK(a) of carboxylic acids using the ab initio continuum-solvation model PCM-UAHF**. *Journal Of Physical Chemistry A*, **102**(33):6706–6712, 1998. 40

[219] Hwang S., Jang Y. H., Cho H., and Lee Y. J. **Density functional theoretical study on the proton migration in radical cations of substituted cytosine : guanine pair**. *Bulletin of the Korean Chemical Society*, **29**(3):539–540, 2008. Times Cited: 0. 41

[220] Camaioni D. M. and Schwerdtfeger C. A. **Comment on "Accurate experimental values for the free energies of hydration of H+, OH-, and H3O+"**. *Journal of Physical Chemistry A*, **109**(47):10795–10797, 2005. 43, 44

[221] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. **Gaussian, Inc., Wallingford, CT**. *Gaussian 03 revision C.02*, 2006. 44

[222] Becke A.D. **Density functional exchange energy approximation with correct asymptotic behavior**. *Physics Review A*, **38**:3098–3100, 1988. 44

[223] Becke A.D. **Density functional theory. III. The role of exact exchange.** *Journal Of Chemical Physics*, **98**:5648–5652, 1993. 44

[224] C. Lee and W. R. G. Yang. **37**. *Phys. Rev. B*, **785**, 1988. 44

[225] Hehre W.J, Ditchfield R., and Pople J. A. **Selfconsistent molecular orbital methods. XII. Further extensions of Gaussiantype basis sets for use in molecular Orbital studies of organic molecules**. *J. Chem. Phys.*, **56**:2257–2261, 1972. 44

[226] Hariharan P.C. and Pople J. A. **The influence of polarization functions on molecular orbital hydrogenation energies**. *Theoret. Chim. Acta*, **28**:213–222, 1973. 44

[227] Hariharan P.C. and Pople J. A. **Accuracy of AH, equilibrium geometries by single determinant molecular orbital theory**. *Molecular Physics*, **27**:209–214, 1974. 44

[228] Gordon M. S. **The isomers of silacyclopropane**. *Chemical Physics Letters*, **76**:163–168, 1980. 44

[229] Francl M. M., Pietro W. J., Hehre W. J., Binkley J. S., Gordon M. S., Defrees D. J., and Pople J. A. **Self-consistent molecular-orbital methods .23. A polarization-type basis set for 2nd-row elements**. *Journal Of Chemical Physics*, **77**(7):3654–3665, 1982. 44

[230] Cossi M., Barone V., Mennucci B., and Tomasi J. **Ab initio study of ionic solutions by a polarizable continuum dielectric model**. *Chemical Physics Letters*, **286**(3-4):253–260, 1998. 45

[231] Mennucci B. and Tomasi J. **Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries**. *Journal Of Chemical Physics*, **106**(12):5151–5158, 1997. 45

[232] Tomasi J., Mennucci B., and Cammi R. **Quantum mechanical continuum solvation models**. *Chemical Reviews*, **105**(8):2999–3093, 2005. 45, 78

[233] Rappe A. K., Casewit C. J., Colwell K. S., Goddard W. A., and Skiff W. M. **UFF, A FULL PERIODIC-TABLE FORCE-FIELD FOR MOLECULAR MECHANICS AND MOLECULAR-DYNAMICS SIMULATIONS**. *Journal of the American Chemical Society*, **114**(25):10024–10035, 1992. Times Cited: 1515. 45, 78

[234] Kendall R. A., Dunning T. H., and Harrison R. J. **Electron-Affinities of the 1st-Row Atoms Revisited - Systematic Basis-Sets and Wave-Functions**. *Journal of Chemical Physics*, **96**(9):6796–6806, 1992. 45

[235] A. Bacchi, M. Balordi, R. Cammi, L. Elviri, C. Pelizzi, F. Picchioni, V. Verdolino, K. Goubitz, R. Peschar, and P. Pelagatti. **4462**. *Eur. J. Inorg. Chem.*, 2008. 45, 64

[236] Scott A. P. and Radom L. **Harmonic vibrational frequencies: An evaluation of Hartree- Fock, Moller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors**. *Journal Of Physical Chemistry*, **100**(41):16502–16513, 1996. OCT 10 J PHYS CHEM. 45

[237] McQuarrie D. A. *Statistical Thermodynamics*. University Science Books, Mill Valley, CA, 1973. 45

[238] Ben-Naim A. and Marcus. **81**. *J. Chem. Phys.*, **2016**, 1984. 45

[239] H. Nakatsuji. **96**. *J. Am. Chem. Soc.*, **24**, 1974. 55

[240] C.M. Zicovich-Wilson, A. Corma, and P. Viruela. **98**. *J. Phys. Chem.*, **10863**, 1994. 55

[241] K.B. Wiberg, P.R. Rable, and P.R. Marquez. **114**. *J. Am. Chem. Soc.*, **8654**, 1992. 56

[242] M. Citroni, M. Ceppatelli, R. Bini, and V. Schettino. **367**. *Chem. Phys. Lett.*, **186**, 2003. 56

[243] E. Whalley. **50**. *Rev. Phys. Chem. Jpn.*, **119**, 1980. 56

[244] M.P. Dare-Edwards, D.J. Gardiner, and N.A. Walker. **316**. *Nature*, **616**, 1985. 56

[245] Semmelhack M. F. **295**. *Ann NY Acad Sci*, **36**, 1977. 68

[246] M. Delferro, M. Tegoni, V. Verdolino, D. Cauzzi, C. Graiff, and A. Tiripicchio. **Oxidative Addition of Iodomethaneto Charge-tuned RhI Complexes**. *Organometallics*, **revision**, 2009. 72

[247] M. P. Andersson and P. Uvdal. **109**. *J. Phys. Chem. A*, **12**, 2005. 76

[248] Fasman G.D. *CRC Handbook of Biochemistry and Molecular Biology, Nucleic Acids*, **1**. CRC Press, Cleveland, OH, 3rd edition edition, 1975. 78, 81, 82, 83, 84

[249] Vincenzo Verdolino, Roberto Cammi, Barbara H. Munk, and H. Bernhard Schlegel. **Calculation of pKa Values of Nucleobases and the Guanine Oxidation Products Guanidinohydantoin and Spiroiminodihydantoin using Density Functional Theory and a Polarizable Continuum Model**. *Journal of Physical Chemistry B*, **in pub.**, 2008. 78, 79, 82, 86, 90, 93, 95, 100, 101, 105

[250] Lide D.R. *CRC Handbook of Chemistry and Physics*, **80th Edition**. CRC Press, Boca Raton, 1999-2000. 81, 82, 84

[251] Messina S. A. and Dawson R. **Attenuation of oxidative damage to DNA by taurine and taurine analogs**. In *Taurine 4: Taurine and Excitable Tissues*, **483** of *Advances in Experimental Medicine and Biology*, pages 355–367. unknown, 2000. 81, 82, 83, 84

[252] Ts'o P.O.P. *Basic Principles in Nucleic Acid Chemistry*. Academic Press, New York, 1974. 81, 82, 83, 84

[253] Jordan D.O. *The Chemistry of Nucleic Acids*. Butterworth and Co., Washington, 1960. 81, 82, 83, 84

[254] Davidson J.N. Chargaff E. *The Nucleic Acids Chemistry and Biology*. Academic Press, New York, 1955. 81, 82, 83, 84

REFERENCES

[255] Bundari S. *The Merck Index*. Merck and Company, Whitehouse Station, NJ, 12th edition edition, 1996. 81, 82, 83, 84

[256] Dawson R.M.C.and Elliott D.C., Elliott W.H., and Jones K.M. *Data for Biochemical Research*. Oxford University Press, Oxford, 3rd edition edition, 1986. 81

[257] Cho B. P. **STRUCTURE OF OXIDATIVELY DAMAGED NUCLEIC-ACID ADDUCTS - PH-DEPENDENCE OF THE C-13 NMR-SPECTRA OF 8-OXOGUANOSINE AND 8-OXOADENOSINE**. *Magnetic Resonance in Chemistry*, **31**(12):1048–1053, 1993. 81, 84, 85

[258] Bausch M., Selmarten D., Gostowski R., and Dobrowolski P. **4**. *J. Phys. Org. Chem.*, **67**, 1991. 81, 89

[259] J. Albert A. Goldacre R. Phillips. **The Strength of Heterocyclic Bases**. *Journal of the Chemical Society*, page 2240, 1948. 81, 86

[260] Dean J. A. *Lange's Handbook of Chemistry*. McGraw Hill, 1999. 81

[261] Shugar David and Fox Jack J. **Spectrophotometric studies op nucleic acid derivatives and related compounds as a function of pH : I. Pyrimidines**. *Biochimica et Biophysica Acta*, **9**:199–218, 1952. 83, 84

[262] Sowers L. C., Shaw B. R., Veigl M. L., and Sedwick W. D. **177**. *Mutat. Res.*, **201**, 1987. 83

[263] Kleinpeter E., Heydenreich M., Kalder L., Koch A., Henning D., Kempter G., Benassi R., and Taddei F. **111**. *J. Mol. Struct.*, **403**, 1997. 86, 95, 100

[264] Kleinpeter E. **161**. *Struct. Chem.*, **8**, 1997. 86

[265] Cristiani F., Devillanova F. A., Diaz A., Isaia F., and Verani G. **487**. *Spectroc. Acta Pt. A-Molec. Biomolec. Spectr.*, **41**, 1985. 86

[266] Sohar P. **165**. *Acta Chim. Sci. Hung.*, **57**, 1986. 86

[267] Sohar P., Nyitrai J., Zauer K., and Lempert K. **189**. *Acta Chim. Sci. Hung.*, **65**, 1970. 86, 95

[268] ACS. **n.d.** *Sci Finder Scholar Registry American Chemical Society*, **n.d.**, 2008. 93, 100, 105

[269] de Oliveira S. M., da Silva J. B. P., Hernandes M. Z., de Lima M. D. A., Galdino S. L., and Pitta I. D. **614**. *Quim. Nova.*, **31**, 2008. 100

[270] Fujioka H. and Tan T. **475**. *J. Pharm. Bio-Dyn.*, **5**, 1982. 100

[271] Anslyn E. V. and Dougherty D. A. *Modern Physical Organic Chemistr*. University Science Books, Sausalito CA, nd edition, 2006. 101

[272] R. Cammi, B. Mennucci, and J. Tomasi. **104**. *J. Phys. Chem. A*, **4690**, 2000. 107

[273] W. Jaskolski. **271**. *Phys. Rep.*, **1**, 1996. 108

[274] R. Noyori and S. Hashiguchi. **30**. *Acc. Chem. Res.*, **97**, 1997. 108

[275] I. M. Pastor and H. Adolfsson. **2046**. *Chem. Commun.*, 2002. 108

[276] A. Bgevig, I. M. Pastor, and H. Adolfsson. **10**. *Chem. Eur. J.*, **294**, 2004. 108