Computer-Aided Drug Design of

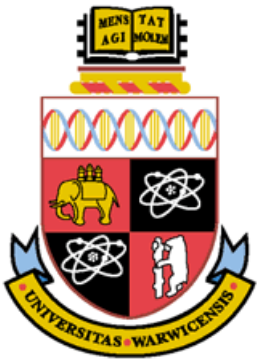
Photoactivated Platinum Anticancer Complexes

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Chemistry

by

Hui-Chung Tai



University of Warwick, Department of Chemistry

October 2011

Table of Contents

Acknowledgements

"You must immerse yourself in an unfamiliar world in order to truly understand your own". For that, I am grateful to the wonderful people with whom I have had the pleasure to share my joys and sorrows of everyday life during my three years in England. If I have built any character or broaden my horizons in scientific and non-scientific aspects, it is all because of you! The friendship and passion that we have cultivated throughout moments together shall never fade away in my life.

First and foremost, I would like to thank **Professor Robert J. Deeth** for his tuition, guidance, and education on Monty Python. Rumour has it that undergraduates often gaze him in admiration as behind the towering facade as a professor in cutting-edge computational chemistry, he has aerobatics under his belt and a great sense of humour. I am profoundly indebted to **Professor Peter J. Sadler** for offering me this great opportunity to engage in the pioneering research in a multicultural team with various expertises. I am grateful for all their supervision, encouragement, and liberal attitude toward science and education.

Many thanks to the many colleagues I have worked with and grown with in **PJS and RJD groups** over the course of the research. In particular, I would like to thank my fellow group members, **Dr. Ralf Brodbeck** for his tuition in computer programming, the help with the simulations, stimulating discussions about thoretical chemistry, and suggestions for many useful books and journal articles, **Mr. Yao Zhao** for all his help with the synthesis, characterisation, and measurement of the spectra for the computed platinum complex as well as **Dr. Anna Anastasi**, **Dr. Christian Diedrich**, **Dr. Nicky Farrer**, **Dr. Ana Pizarro**, **Dr. Luca Salassa**, and **Dr. Hazel Phillips** for helpful discussions.

I gratefully acknowledge the invaluable comments and advices from **Professor P. Mark Rodger** and **Professor Gregory L. Challis** in my advisory panel during my Ph.D. study. In addition, I would like to thank our IT manager, **Mr. Jason Noone**, for all his help with the maintenance of software and hardware in Warwick Chemistry.

Special thanks to **Dr. Abraha Habtemariam** for imparting his knowledge, insights, and skills in pool in countless interesting conversions over cup of tea and in even more exciting chats over glass of beer. A huge thank you must go to **Miss** (Mrs.-to-be) **Claire Booyjzsen** for being my best friend and pro bono English teacher. Her sincerity, patience, and diverse artistic pursuits and social bonding have added warmth and colour to my study abroad. I simply cannot find the words to thank her enough for those marvelous moments we have shared. Thanks also to **Dr. Sarah Farley** for all her help at the beginning of my study. I would like to thank **Dr. Julie Ann Lough** for her generosity, friendship and one of a kind personality. For me, she is truly 'Dr. Sans Frontieres'. I am truly grateful to **Dr. Miho Wako** for being an attentive flatmate. Those days we spent with each other has been stamped indelibly on my memory. Sincere thanks to **Dr. Yvonne Huang** for her witty and boozy conversations in my never-ending process of writing thesis.

I would also like to thank the **University of Warwick** for financial support during my study. My graduate extends to all the staff in the **Research Student Skills Programme.** Thanks to them, my work has become more and more contagious among audiences.

In the voyage of becoming a decent scientist, I appreciate **Dr. Dirk V. Deubel** being my mentor and introducing me to computational chemistry. Without his help and encouragement, I would not be able to overcome the adversity toward my very enjoyable Ph.D. study in University of Warwick. Special thanks to **Professor Carmay Lim** for helping me develop critical thinking while working on Master thesis and under the daunting interrogation in group meeting and oral examination. In addition, I would like to thank **Professor Yuan-Pern Lee** for his illuminating and creative lectures in physical chemistry at National Tsing-Hua University which is the fundamentals of many areas of research. His encouragement and enthusiasm for science and teaching have a forever inspiration in my navigation through the ocean of the unknown.

Finally, I wish to acknowledge the support of all my dear friends and family in Taiwan. I am extremely grateful to my **Mother** for all her dedication and inspiration. Without her endless love, I can never imagine myself coming this far. I would like to express my gratitude to **Dr. Yen-ling Lee** for being my best friends since our adolescence and always being there for me. Thanks also to Dr**. Yao Chi Chen** for all the immediate help in my key steps of becoming a scientist and her warm friendship.

Declaration

I hereby declare that the work in this thesis is the original work of the author, except where acknowledgement has been made to other sources. The work was undertaken at the Department of Chemistry, University of Warwick between January 2008 and December 2010 and has not been submitted, in whole or in part, for any other degree, diploma or other qualification. A list of research papers in preparation from this work is given below.

Hui-Chung Tai

October 2011

Publications

1. **Synthesis, Characterisation and Photochemistry of Pt-IV Pyridyl Azido Acetato Complexes**
   Mackay, F. S.; Farrer, N. J.; Salassa, L.; Tai, H-C.; Deeth, R. J.; Moggach, S. A.; Wood, P. A.; Parsons, S.; Sadler, P. J.
   Dalton Trans. **2009**, (13), 2315.

2. **A Computational Approach to Tuning the Photochemistry of Pt(IV) Anticancer Agents**
   Tai, H-C.; Zhao, Y.; Farrer, N. J.; Anastasi, A. E.; Deeth, R. J.; Sadler P. J.
   Chem. Eur. J. **2012**, *18* (34),10630.

3. **Combined Theoretical and Computational Study of Interstrand DNA Guanine-guanine Cross-linking by *trans*-[Pt(pyridine)$_2$] Derived from the Photoactivated Prodrug *trans,trans,trans*-[Pt(N$_3$)$_2$(OH)$_2$(pyridine)$_2$]**
   Tai, H-C.; Brodbeck, R.; Jana Kasparkova, Nicola J. Farrer, Viktor Brabec, Sadler, P. J.; Deeth, R. J.
   Inorg. Chem. **2012**, *51*(12), 6830.

Abbreviations

AOM	Angular Overlap Model
CFT	Crystal Field Theory
CSD	Cambridge Structure Database
DFT	Density Functional Theory
DommiMOE	d orbital molecular mechanics in Molecular Operating Environment
FDA	Food and Drug Administration
5FU	5-fluorouracil
FWHM	Full Width at Half Maximum
HMG	High Mobility Group
ISC	Intersystem crossing
LFMD	Ligand Field Molecular Dynamics
LFMM	Ligand Field Molecular Mechanics
LFO	Ligand Field Orbital
LFSE	Ligand Field Stabilisation Energy
LFT	Ligand Field Theory
LLCT	Ligand-to-Ligand Charge Transfer
LMCT	Ligand-to-Metal Charge Transfer
MLCT	Metal-to-Ligand Charge Transfer
MM	Molecular Mechanics
MO	Molecular Orbital
PDB	Protein Data Bank
PDT	Photodynamic Therapy
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
RMSD	Root-mean-square Deviation
SAOP	Statistical Average of Orbital Potentials
TBP	TATA-binding protein
TDDFT	Time-dependent Density Functional Theory
ZORA	Zeroth-Order Regular Approximation

Abstract

Platinum(IV) complexes are usually inert and stable compounds which can be photoactive pro-drugs to produce Pt(II) species with promising anti-cancer activity. Studies of the photochemistry of Pt(IV) complexes by time-dependent density functional theory (TDDFT) and spectroscopic methods show close agreement. Broad exploration of *cis*/*trans* geometries, *trans* influences, the nature of the OR^- and (pseudo)halogen ligands, electron-withdrawing/donating/ delocalizing substituents on the N-ligands, and intramolecular H-bonds shows that (1) the design of platinum(IV) complexes with intense bands shifted towards longer wavelengths (~330 nm) can be achieved by introducing intramolecular H-bonds involving the OH ligands and 2-hydroxyquinoline or by iodido ligands, (2) mesomeric electron-withdrawing substituents on pyridine result in low-energy absorption with significant intensity in the visible region, and (3) the distinct makeup of the molecular orbitals in electronic transitions for *cis*/*trans*-{$Pt(N_3)_2$} isomers result in different photoproducts. In general, the comparison of the optimised geometries shows that Pt(IV) complexes with longer Pt−L bonds are more likely to undergo photoreduction with longer-wavelength light. Complex *trans*, *trans*, *trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] was first synthesised. The experimental UV-Vis spectrum in aqueous solution correlates well with the intense band in the computed spectrum whereas the overlay in the low-energy absorption region can be improved by a solvent model. This combined computational and experimental study shows that TDDFT can be a design tool to tune the coordination environment for optimizing photoactive Pt(IV) compounds as anticancer agents without immediate need for synthesis.

Additionally, molecular modeling is used to study DNA distortions induced by binding metal-containing fragments derived from *cis*-diamminodichloroplatinum(II) (cisplatin) and a new class of photoactive platinum anticancer drugs. Ligand field molecular mechanics (LFMM) parameters for Pt–guanine interactions are derived and validated against a range of experimental structures from the Cambridge Structural Database, published quantum mechanics/molecular mechanics (QM/MM) structures of model Pt-DNA systems and additional DFT studies. LFMM gives a good description of the local Pt-guanine coordination at a fraction of the computational cost of QM/MM methods. The force field is then used to develop protocols for ligand field molecular dynamics (LFMD) simulations using experimentally characterised bifunctional DNA adducts involving both an intra- and an interstrand crosslink of cisplatin as a prelude to studying the interaction of *trans*-$\{Pt(py)_2\}^{2+}$ (**P**, py = pyridine), the major photoproduct of the novel platinum(IV) complex *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ (**17**), with the DNA duplex dodecamer, $d(C_1C_2T_3C_4T_5C_6G_7T_8C_9T_{10}C_{11}C_{12})\bullet\ d(G_{13}G_{14}A_{15}G_{16}A_{17}C_{18}G_{19}A_{20}G_{21}A_{22}G_{23}G_{24})$. Based on the observed formation of a *trans* species when **17** is photoreduced in the presence of 5'-guanosine monophosphate plus the major-groove binding mode of the monofunctional complex *cis*-$\{Pt(NH_3)_2(py)\}^{2+}$, **P** is proposed to coordinate to G_7 and G_{19}. This **P**-DNA adduct has a widened minor groove at one end of the platinated site and deepened minor groove at the opposite end, and exhibits a global bend of ~67° and an unwinding of ~20°. Brabec *et al.* subsequently demonstrated experimentally that such interstrand GG crosslinks can form as a result of the photoactivation of **17** in the presence of DNA. Such cross-links offer

possibilities for specific protein–DNA interactions and suggest possible mechanisms to explain the high potency of this photoactivated complex.

Chapter 1

Introduction

“All substances are poisonous, there is none that is not a poison; the right dose differentiates a poison from a remedy.” (Von der Besucht Paracelsus, 1567)

The overall scope of this thesis is the implementation and application of computational methods for understanding the structures, mechanisms of photoactivation, and modes of action of photoactivated anticancer platinum complexes to further aid drug design and development. In recent years, the significant increase of molecular modeling as an important integral part of multidisciplinary efforts in the design and synthesis of new potent pharmaceuticals has spurred the interest of scientists with various backgrounds on its application. Therefore, in addition to the contribution of computational approaches in rational drug design, this thesis has attempted to cover the current status of platinum-based drug design, the theoretical methods applicable to the research, and the required information for calculating and interpreting the experimentally observed properties on the basis of the theoretical results in the hope of serving as a useful reference for anyone who works in the multidisciplinary field of drug design and discovery and shows a zest for computer-aided drug design of metallopharmaceuticals.

Chapter 1 presents a collection of overviews of platinum-based anticancer drugs, strategies in drug design and development, fundamental photochemical principles, and the design of photoactive platinum complexes as prodrugs. Chapter 2 provides an introduction to computational methods for modelling and designing molecules. Chapters 3, 4, and 5 cover my novel contributions. Chapter 3 presents a computational approach to tuning the photochemistry of Pt(IV) anticancer agents. In Chapter 4, the first application of ligand field molecular mechanics (LFMM) for modelling of platinum–DNA complexes is described. Chapter 5 focuses on the application of the novel LFMM parameters developed in Chapter 4 in simulations of a variety of platinum–DNA cross-links. The final Chapter 6 presents a

perspective on drug design and suggests future work to advance the rational design of metal-based anticancer complexes.

1.1 Platinum-Based Anticancer Drugs

The discovery of the anticancer activity of cisplatin (*cis*-$[PtCl_2(NH_3)_2]$) and its clinical introduction in the 1970s is an important landmark in the history of successful treatment of cancers. Despite the introduction in the 1980s of carboplatin, a second-generation analogue that is safer but shows a similar spectrum of activity to cisplatin, the pace of further improvements slowed for decades. However, recent key advances in the elucidation of the mechanisms of action of the drugs and of tumour resistance to the drugs, the approval of new platinum-based agents (*e.g.* oxaliplatin), and clinical studies using platinum drugs in combination with resistance modulators or new molecularly-targeted drugs, have lead to a surge in interest in developing new anticancer platinum drugs.[1]

1.1.1 Serendipitous Discovery

The interest in platinum-based antitumor drugs has its origin in the 1960s, with the accidental discovery by Rosenberg of the inhibition of cell division by Pt complexes. Detailed chemical analysis identified two active complexes — *trans*-diamminetetrachloroplatinum(IV) (*trans*-$[Pt(NH_3)_2Cl_4]$) and *cis*-diamminetetrachloroplatinum(IV) (*cis*-$[Pt(NH_3)_2Cl_4]$) — as the molecules causing filamentation due to cell growth without proper division. The *trans* isomer was much less active. The active platinum(II) analogue — *cis*-$[Pt(NH_3)_2Cl_2]$, which went on to be cisplatin, was made en route to synthesising the Pt(IV) complexes, and it also was active in forming filaments.[2] As it turned out, the group had rediscovered a known platinum coordination complex, known as Peyrone's

chloride, that was originally synthesised and described in 1845.[3] Followed by confirmatory *in vivo* tests in London, cisplatin underwent clinical testing in US. A very short time after the original 'bench' discovery, the first patients were treated in 1971. Approval of cisplatin for the treatment of testicular and ovarian cancer by the US Food and Drug Administration (FDA) was granted in 1978. Today, cisplatin is one of the most widely utilised antitumor drugs in the world. Its discovery initiated a new chapter of research in medicinal inorganic chemistry and has led to the synthesis and biological testing of over 3000 cisplatin analogues,[4] and study on other nearby elements from the periodic table (e.g. ruthenium, palladium, and gold). Despite its highly effective treatment in testicular and ovarian cancers, cisplatin has several disadvantages that include severe toxicity such as nephrotoxicity, neurotoxicity, and emetogensis.[5] Tremendous efforts in the design of new platinum drugs were aimed at making chemotherapy safer to patients. The development of new platinum-based drugs has encountered many twists and turns over the past three decades (see timeline in Figure 1-1). At present, only four platinum complexes are registered as marketed drugs (cisplatin, carboplatin, oxaliplatin, and nedaplatin in Table 1-1.)[6]

1.1.2 How Does Cisplatin Work?

Upon administration, the relatively high concentration of chloride ions (~100 mM) maintains cisplatin in a neutral state. Inside the cell, the lower ambient chloride ion concentrations (~4-20 mM) facilitates the activation of cisplatin by stepwise aquation, and subsequent reaction with various cellular targets, most notably

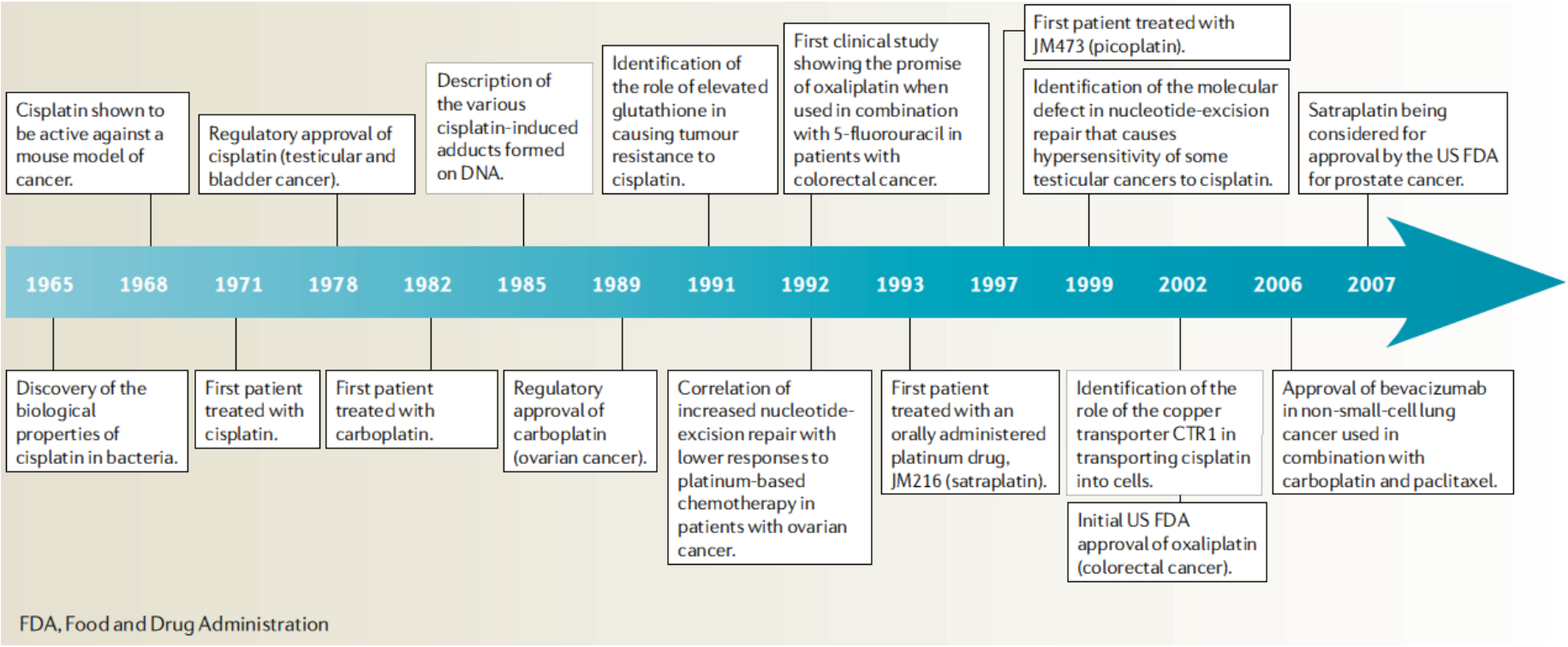
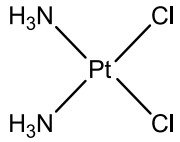
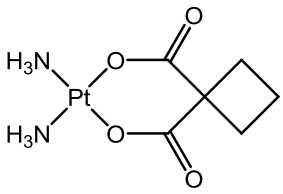
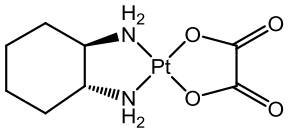
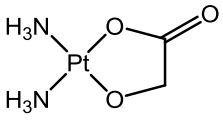


Figure 1-1. Milestones in the development of platinum drugs for cancer therapy. Reproduced from reference 1 with permission from Nature Publishing Group.

5

Table 1-1. Marketed platinum drugs (5FU = 5-fluorouracil; LV = leucovorin). Adapted from reference 1,6.

Drug	Approval year		Dose-limiting toxicity
 Cisplatin	Metastatic testicular cancer Metastatic ovarian cancer Transitional bladder cancer	1978 1978 1993	Nephrotoxicity Neurotoxicity, ototoxicity Nausea and vomiting
 Carboplatin	Ovarian cancer (palliative after previous chemotherapy) Ovarian cancer, first line	1989 1991	Myelosuppression (thrombocytopenia and neutropenia) Nausea and vomiting (but less than with cisplatin)
 Oxaliplatin	Accelerated approval, metastatic colorectal cancer (second line with 5FU with LV) Colorectal cancer (previously untreated or adjuvant treatment with 5FU with LV)	2002 2004	Neurotoxicity (sensory peripheral neuropathy) Nausea and vomiting
 Nedaplatin	Approval in Japan, cervical cancer, and non-small and small cell lung cancer	1995	myelosuppresion

DNA. The interaction between platinated DNA and nuclear proteins induces signal transduction pathways which lead to DNA-damage recognition and repair, cell-cycle arrest, and apoptosis or necrosis.[6,7] Experimental characterisation of the types of platinum–DNA adducts formed shows the presence of approximately 65% 1,2-d(GpG), 25% 1,2-d(ApG), and 5-10% 1,3-d(GpNpG) intrastrand cross-links as major components, although interstrand cross-links and monofunctional adducts have also been identified (Figure 1-2).[6] In all cases, platinum is bound to the N7 position of purine bases (Figure 1-2). Formation of cisplatin adducts significantly alters the structure of the target DNA by bending, unwinding, and

destabilising the duplex. The structural details of platinum-DNA adducts are available from X-ray, NMR and molecular mechanics simulation.[8-17] Although these DNA adducts formed by platinum drugs exhibit some degree of structural similarity, arising from coordination to the N7 atom of the purine base, it is clear that each distorts the duplex DNA in a distinctive manner.[11,18] These nuances may activate distinctive DNA-damage recognition and cellular processes, which possibly mediate cytotoxicity and anticancer properties of the compounds via different mechanisms of action.[1,6,7]

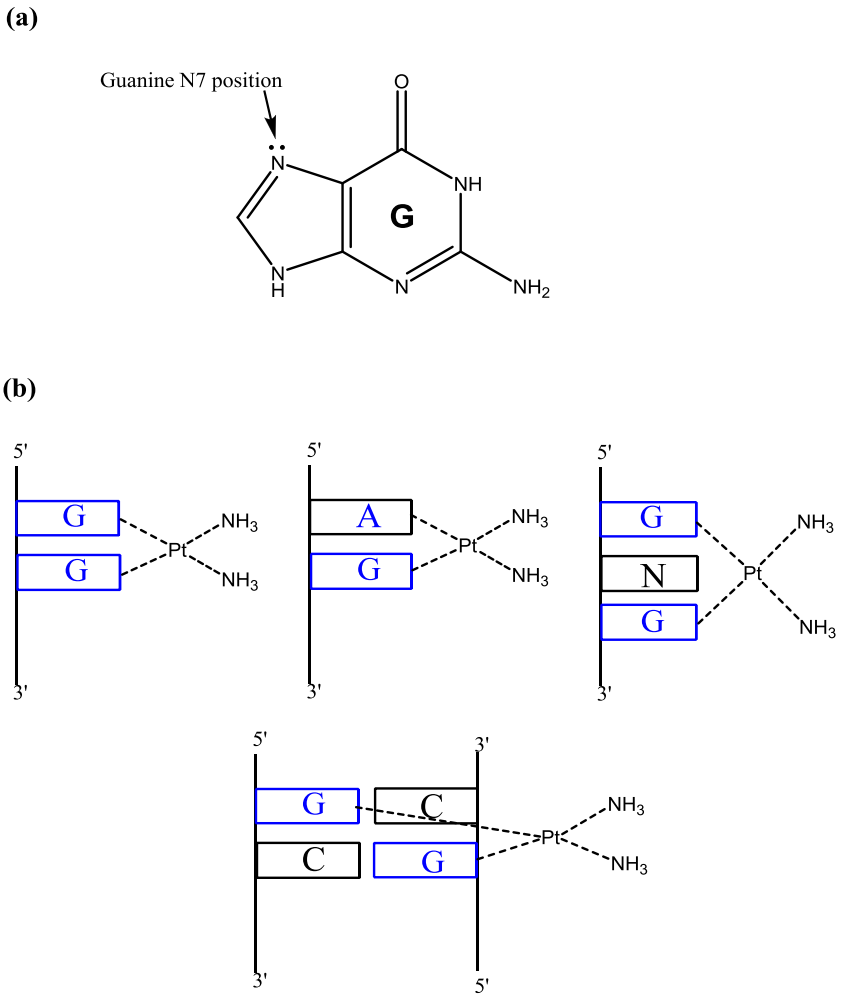


Figure 1-2. Platinum preferential binding at N7 sites of purine bases and bifunctional cisplatin-DNA adducts: 1,2-d(GpG), 1,2-d(ApG), and 1,3-d(GpNpG) intrastrand as well as interstrand cross-links.

Platinum-based drugs disrupt DNA synthesis by mechanisms common to all actively dividing cells, regardless of whether they are cancerous or not, which generates severe side effects. In addition, the large volume of damaged cells further cause altered cellular transport, enhanced repair of distorted DNA, overexpression of DNA-damage recognition proteins and decreased apoptosis contributing to tumour resistance to platium drugs. [6,19] Thus, a thorough understanding of the mechanism of action of platinum-based anticancer agents and the design of drugs with increased selectivity are of great importance for significant clinical advantages over the current drugs.

1.2 Rational Design of Platinum Anticancer Drugs

1.2.1 Structure-Activity Relationship

All Pt compounds that have entered clinical trials so far adhere to the structure-activity relationships summarized by Cleare and Hoeschele.[20] These relationships state that for a Pt complex to show antitumor activity, the Pt(II) or Pt(IV) complex should have a *cis* geometry with the general formulas of *cis*-$[PtX_2(Am)_2]$ or *cis*-$[PtX_2Y_2(Am)_2]$, where X is the leaving group and Am is an inert amine with at least one N–H moiety. The leaving group, X, should be an anion with intermediate binding strength to platinum and have a weak *trans* effect to avoid labilizing the amine. Complexes with labile leaving groups such as ClO_4^- or NO_3^- are highly toxic, while complexes with inert leaving groups are generally inactive. The guidelines have dominated platinum drug design for decades and remained valid until relatively recently. However, it has become quite evident that mere analogues of cisplatin or carboplatin will probably not offer any substantial clinical advantages over the existing drugs. A number of compounds that violate

the structure-activity relationships, yet show antitumor activities, have been discovered. Efforts have also been directed toward the rational design of compounds with specific characteristics that make chemotherapy safer for patients, in particular, lessening severe side effects, increasing oral bioavailability, and overcoming drug resistance.[5]

1.2.2 Strategies in the Development of New Platinum Anticancer Agents

Herein, the generalized rationale underlying the development of the current platinum anticancer drugs is summarised.[1] Modification of the structure of the leaving groups (e.g. carboplatin) appears to alter the physical and pharmacokinetic properties and thus influence tissue and cellular distribution of the platinum coordination complexes, but it is unlikely to modify pharmacological properties. In contrast, modification of the amine carrier ligands (e.g. oxaliplatin) could lead to different pharmacokinetics and structurally different DNA lesions possibly associated with a different spectrum of activity. Substitution of the more stable cyclobutanedicarboxylate for the two chlorides results in less toxic carboplatin, while replacement of the two ammonia ligands by the sterically-hindered diaminocyclohexane leads to oxaliplatin, which retains activity against some cancer cells with acquired resistance to cisplatin.[21] Of the various strategies that have been developed to enhance selective targeting of a tumour-killing agent to tumours, *prodrugs* that can deliver platinum in a relatively inactive form and subsequently be released and activated at the tumour site appear to have particular potential. Improved delivery of platinum drugs with liposomal-based or polymer-based vehicles is an attractive strategy and has been in clinical trials.[22-24] However, creating delivery systems that target specific sites to achieve good release and

activation is challenging. On the other hand, a photoactive agent in combination with light provides a promising avenue to achieve accurate targeting.[25,26]

1.3 Photochemical Principles

Light is a form of energy that exhibits both wave-like and particle-like properties. The intensity of light is associated with the number of photons per unit time; the energy (E) of the photons is related to the frequency (v) of the radiation by the expression:

$$E = h \times v \tag{1-1}$$

where h is Planck's constant. An increase in the intensity of radiation will therefore only increase the number of species excited and not the energy available to each individual species. Energy levels are quantized; i.e. species can exist only in certain defined, discrete energy states. A direct result of the quantization of energy levels is that for each individual species, only specific energies and, therefore, frequencies of radiation can be absorbed or emitted. The characteristic line and band spectra of chemical species are a consequence of this behaviour. When a species absorbs a quantum of radiation, it becomes excited. How the species assimilates that energy — in rotational, vibrational, or electronic modes — depends on the wavelength of the incident radiation. As the differences in quantized energy levels for vibrational energy and those for rotational energy are respectively smaller by nearly 10^2 and 10^4 times than those for the electronic energy, absorption at low energies associated with infra-red radiation usually only leads to transitions between rotational or vibrational states of the species. The electronic transitions require energies in the visible and ultraviolet regions of the

electromagnetic spectrum and are accompanied by simultaneous changes in the vibrational and rotational states.

The excited state of a species can exhibit a dramatically altered reactivity from the ground state. Not only does the species possess more energy, but it can also participate in different reactions as a result of the new electronic arrangement. The energies of electronic excitation are of the same order of magnitude as the bonds in a species. Furthermore, the energies correspond roughly with typical activation energies of reactions, and the excitation energy can help a species to either partially or completely overcome an activation barrier. The arrangement of the electrons in a species is fundamental in defining what reactions the species can and cannot undergo, so that a change in the spatial distribution of electrons can have a major effect on reactivity.[27]

1.3.1 Selection Rules for Optical Transitions

A photon has an intrinsic spin angular momentum corresponding to $s = 1$. The change in angular momentum of the electron must compensate for the angular momentum lost upon the absorption of the photon. Thus, an electron in a *s* orbital ($l = 0$) cannot make a transition into an *d* orbital ($l = 2$) because not enough angular momentum can be lost by the absorption of the photon. Similarly, an *s* electron cannot make a transition to another *s* orbital, because there would then be no change in the electron's angular momentum to make up for the angular momentum lost by the photon. It follows that some spectroscopic transitions are allowed, meaning that they can occur, whereas others are forbidden**,** meaning that they cannot occur. Selection rules are statements about which transitions are allowed and give an indication of the feasibility of a transition from one state to

another in terms of the quantum number. They are derived by identifying the transitions that conserve angular momentum when a photon is emitted or absorbed but give no indication of the absolute intensities of those transitions.

The selection rules can be summarised as follows.

(1) $\Delta \boldsymbol{S} = 0$

(2) In atoms, for one-electron transitions,

$\Delta \boldsymbol{L} = \pm 1$

$\Delta \boldsymbol{J} = 0, \pm 1$, but $\boldsymbol{J} = 0 \nrightarrow \boldsymbol{J} = 0$

(3) For diatomic and linear polyatomic molecules,

$\Delta \boldsymbol{\Lambda} = 0, \pm 1$

The total electronic orbital angular momentum **L** of an n-electron atom is defined as the vector sum of the orbital angular momenta of the individual electrons:

$$\mathbf{L} = \sum_{i=1}^{n} \mathbf{l}_i \qquad (1\text{-}2)$$

The total electronic spin angular momentum **S** of an atom is defined as the vector sum of the spins of the individual electrons:

$$\mathbf{S} = \sum_{i=1}^{n} \mathbf{s}_i \qquad (1\text{-}3)$$

The total electronic angular momentum **J** of an atom is the vector sum of the total electronic orbital and spin angular momenta:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \qquad (1\text{-}4)$$

The value of $\boldsymbol{\Lambda}$ is the sum of the quantum number for the component of the total orbital angular momentum of an individual electron around the internuclear axis. There are two selection rules concerned with changes in symmetry. First, only $+ \leftrightarrow +$ and $- \leftrightarrow -$ transitions are allowed. Second, the Laporte selection rule for molecules and atoms with a centre of inversion states that: the only allowed

transitions are transitions that are accompanied by a change of parity; that is, u ↔ g transitions are allowed, but g ↔ g and u ↔ u transitions are forbidden.[27,28]

1.3.2 Franck-Condon Principle

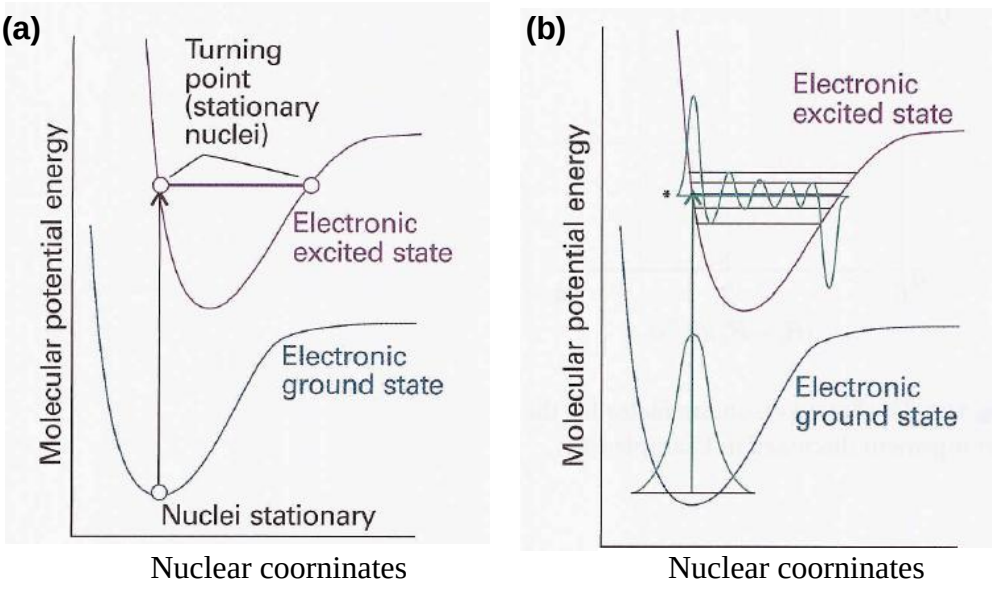


Figure 1-3. (a) The idea of the Franck-Condon principle is that the most intense vibronic transition is from the ground vibrational state to the vibrational state lying vertically above it. Transitions to other vibrational energy levels also occur, but with lower intensity; (b) In the quantum mechanical picture of the Franck-Condon principle, the molecule undergoes a transition to the upper vibrational state with wavefunction that has the greatest overlap integral with the vibrational wavefunction of the vibrational ground state of the lower electronic state. Adapted from reference 28.

Because nuclei are so much heavier than electrons, an electronic transition takes place very much faster than the nuclei can respond. Consequently, electron density is rapidly moved into new regions of the molecule via vertical transition. The initially stationary nuclei suddenly experience a new force field, and in response they begin to swing backwards and forwards from their original positions which were maintained during the rapid electronic excitation. Thus, the

equilibrium nuclear positions in the initial electronic state become stationary turning points in the final electronic state (Figure 1-3a).

Figure 1-3b provides a quantum mechanical view of the Franck-Condon principle. When the molecule is in the lowest vibrational state of its lowest electronic state, the most probable positions of the nuclei are at their equilibrium separation, R_e. The electronic transition is most likely to take place when the nuclei have this separation in the excited states represented by the upper curve, so that the nuclear coordinates undergo the minimal change during the essentially instantaneous electronic transition. We may imagine the transition as being up the vertical line in Figure 1-3a, *vertical transition*, due to the assumption of constant nuclear coordinates during the transition.

The quantum mechanical formulation of the Franck-Condon principle is that the intensity of a vibronic transition is proportional to the square of the overlap integral between the vibrational wavefunctions of the two states involved in the transition. The vibrational wavefunction shows maximum amplitude at the level marked *, which is thus the most probable state for the transition. Other nearby states of the nuclei being at the separation R_e have appreciable probabilities at various degrees. Hence, transitions occur to all the vibrational states in this region, but most intensely to the state with the strongest peaks of vibrational wavefunction near R_e.

The relative horizontal position of the two potential energy curves determines the vibrational structure of the spectrum and the mechanisms of the photoactivated chemical reactions. Electronically excited states usually have more antibonding character than ground states, and thus the upper curve tends to be displaced to longer equilibrium bond lengths.

1.3.3 Photochemistry of Coordination Complexes

For metal complexes, excitation leads to electronically- and vibrationally-excited states with the same multiplicity as the ground state. Figure 1-4 is a schematic localised molecular orbital (MO) diagram for an octahedral transition metal complex. The electronic transitions to the excited states are classified according to the character of the orbitals involved in the electronic transition. This is a simplified classification in some cases. Orbitals may have mixed metal/ligand character depending on the nature of the metal-ligand bond and electronic transitions may involve more than two orbitals at a time. There are principally four types of electronic transitions:

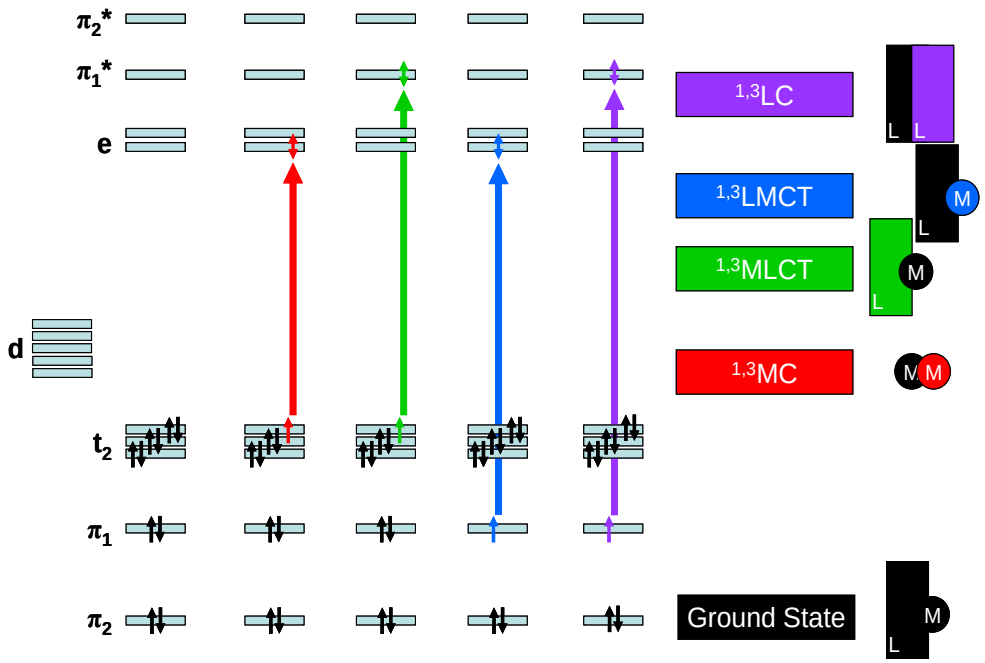


Figure 1-4. Simplified molecular orbital diagram of electronic transitions for a low-spin d^6 metal complex with octahedral coordination (strong crystal field is assumed). Each black arrow (↑ or ↓) represents an electron with its associated spin. Coloured arrows (↕) represent the electron involved in each electronic transition. In the singlet state, excited electrons are spin-down (↓), while in the triplet state they are spin-up (↑). Adapted from reference 29.

Metal-centred (MC) transitions. Electronic transitions occurring between the metal d orbitals are known as *ligand-field* (LF) or *d-d* transitions. These are orbitally (Laporte)-forbidden, and can also be spin-forbidden if the spin state changes. As a result, these transitions are often weak (extinction coefficients ~ 1–1000 M^{-1} cm^{-1} for orbital-forbidden; 0.01–10 M^{-1} cm^{-1} for spin-forbidden) and can be masked by more intense, formally allowed charge-transfer transitions. Since MC transitions typically populate excited states with antibonding character, they often lead to bond lengthening and favour ligand substitution. Photochemical lability is thus commonly a feature of complexes in which a MC excited state is lowest in energy.

Metal-to-ligand charge transfer (MLCT) transitions. Electrons migrate from the metal d orbitals to the ligand orbitals.

Ligand-to-metal charge transfer (LMCT) transitions. Electrons move from orbitals of the ligand to the d obitals of the metal. MLCT and LMCT transitions give rise to stronger absorption (typical extinction coefficients ~ 10–50,000 M^{-1} cm^{-1}) and can bring about redox reactions and also cause homolytic bond cleavage, reducing the metal and producing radicals. Generation of radicals under biological conditions can cause damage to cellular components.

Ligand-centred (LC) transitions. Electron transitions occurring between orbitals of two ligands are known as *interligand* (IL) transitions. These often occur in large delocalised systems.

The excited states can undergo a series of physical radiationless processes including intersystem crossing (ISC), internal conversion (IC) vibrational relaxation, intramolecular vibrational redistribution and solvation dynamics

(reorganisation of solvent shells), which ultimately lead to the ground-state electronic structure. Radiative processes such as fluorescence (singlet-singlet) and phosphorescence (triplet-singlet) result in a return to the ground state with emission of light of longer wavelength than for the excitation. The efficient intersystem crossing promoted by the metal ion prolongs the lifetime of triplet states to the timescale of 50 ns – 1 μs, much faster than that of classical organic compounds (ms), with relatively high emission quantum yield. The main absorption and emission processes can be summarised by a Jablonski diagram (Figure 1-5).[29,30]

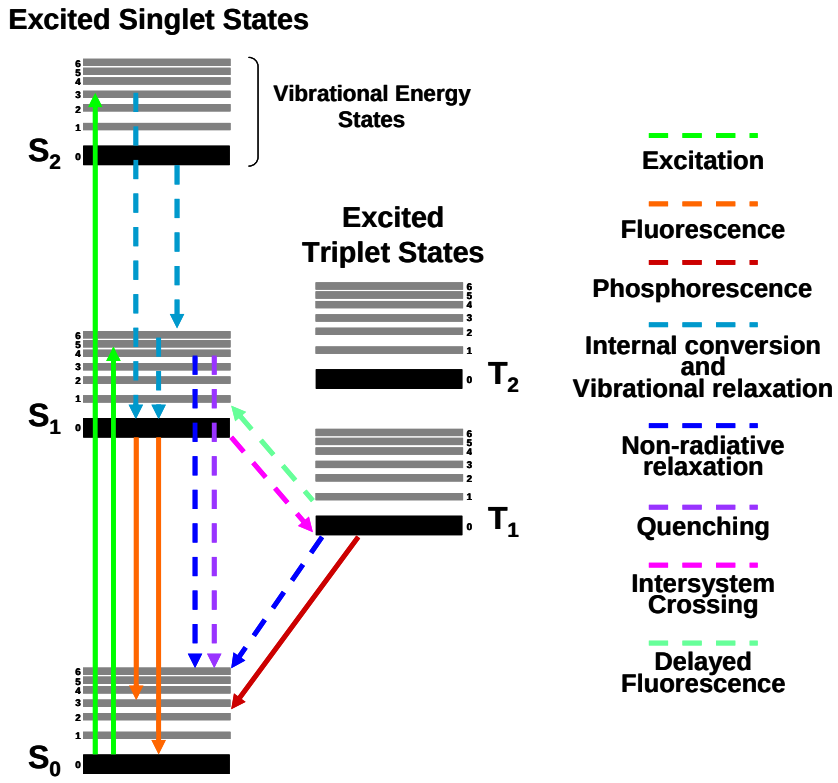


Figure 1-5. Jabłonski energy diagram. The principal physical processes triggered by light excitation of a d^6 metal complex are represented by dashed (radiationless) and solid (radiative) lines. Adapted from reference 29.

1.4 Design of Photoactivated Platinum Anticancer Drugs

The combination of light and metal complexes has been used in medicine. Phototherapy of cancer, i.e. the treatment of a tumour with a drug and light, is particularly attractive as it provides site-specific activation of the *photodrug*, and thus allows for local treatment of the tumour and minimise the side effects of chemotherapy. Recent advances in laser and fiber optic technologies have made it possible to reach specific tissues in the body with light of highly defined intensity and wavelength, extending the applicability of this method beyond easily accessible cancer cells. A key parameter of the approach is the wavelength of the light. For efficient phototherapy, light with wavelengths between 300 and 900 nm is generally preferred, as shorter wavelengths have limited depth of penetration into human tissue and can cause severe damage to tissue, and longer wavelengths lack sufficient energy for chemical transformation.

Photodynamic therapy (PDT) is used clinically to selectively damage easily accessible cancer cells (e.g. from the skin, neck and head, oesophagus, and bladder) by using a photosensitising drug and light. Upon irradiation, the photosensitiser, typically a porphyrin, becomes excited and this energy is transferred to ground state triplet oxygen ($^3O_2$) to create highly reactive singlet oxygen ($^1O_2$) that reacts rapidly with cellular components to cause damage leading to cell death. However, this oxygen-dependent mechanism for cytotoxicity is a major drawback of PDT as many malignant and aggressive cancer cells are hypoxic. It is thus highly desirable to develop photochemotherapy that is not dependent on oxygen.[25,31]

1.4.1 Pt(IV) Prodrugs

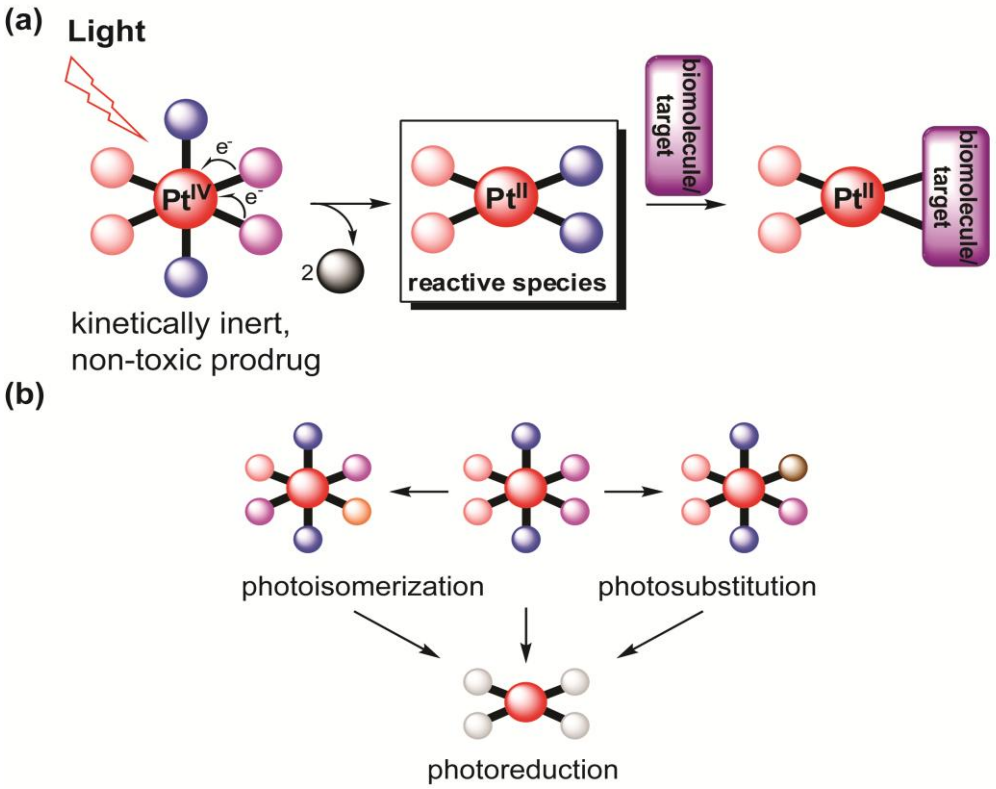


Figure 1-6. Photoactivation of Pt(IV) complexes as a prodrug strategy for metallopharmaceuticals; (a) a general scheme of prodrug activation by photoreduction; (b) photosubstitution and photoisomerization are competing photoreaction pathways, which can generate different reactive species upon reduction. Adapted from reference 32.

The general kinetic inertness of Pt(IV) complexes compared to Pt(II) complexes has been widely exploited in the design of potential prodrugs. The Pt(IV) compound satraplatin with improved oral activity has been abandoned recently in phase III clinical trials for the treatment of hormone-refractory prostate cancer because its effectiveness was not better than current Pt drugs. Substitution reactions of Pt(IV) complexes under physiological conditions are very slow or do not take place at all. Therefore, intracellular reduction to Pt(II) oxidation state is widely accepted to be essential for anticancer activity. This reduction might be achieved by cellular reducing agents or by irradiation with light, enabling site-specific activation of the drugs. However, in addition to photoreduction, the

occurrence of a mixture of photosubstitution and photoisomerization makes the photopathways of Pt complexes difficult to predict and a careful analysis of the photoproducts imperative (see Figure 1-6).[25] Ultimately, the aim is to design Pt(IV) anticancer drugs that specifically target cancer cells, and to have control over ligand-exchange kinetics, and selective activation by light to allow for temporal and spatial control of drug delivery and activation.

1.4.2 Photoactivated Platinum Anticancer Drugs

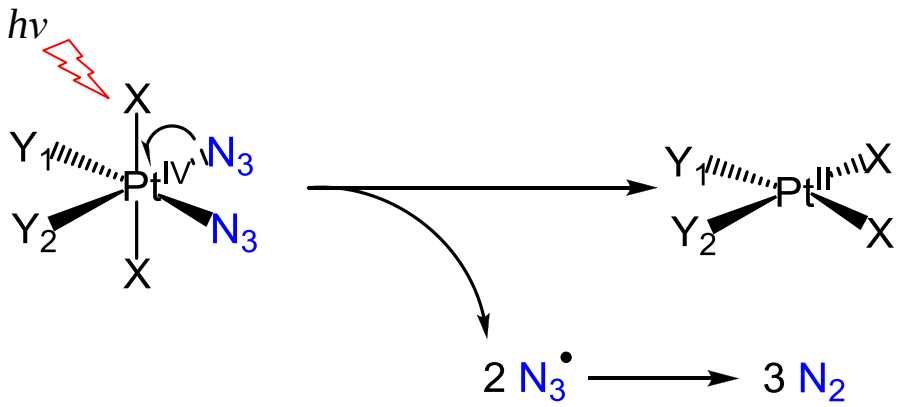


Figure 1-7. Possible mechanism for photoactivation of platinum(IV) diazido complexes.

Trans-dihydroxido platinum(IV) prodrugs containing two azido ligands in *trans* or *cis* positions relative to each other are nontoxic in the dark but show cytotoxicity toward various human cancer cell lines upon photoactivation via a oxygen-independent mechanism.[33,34] In particular, a Pt(IV) diazido complex containing pyridine *trans* to ammonia, *trans,trans,trans*-[$Pt(N_3)_2(OH)_2(py)(NH_3)$], is up to 80 times more cytotoxic than cisplatin in ovarian cancer cells upon irradiation but inactive and stable toward biological reductants in the dark.[35] One photoactivation pathway for these complexes (among several possibilities) involves ligand-to-metal charge transfer from the azido ligands to Pt(IV), resulting in reduction to the

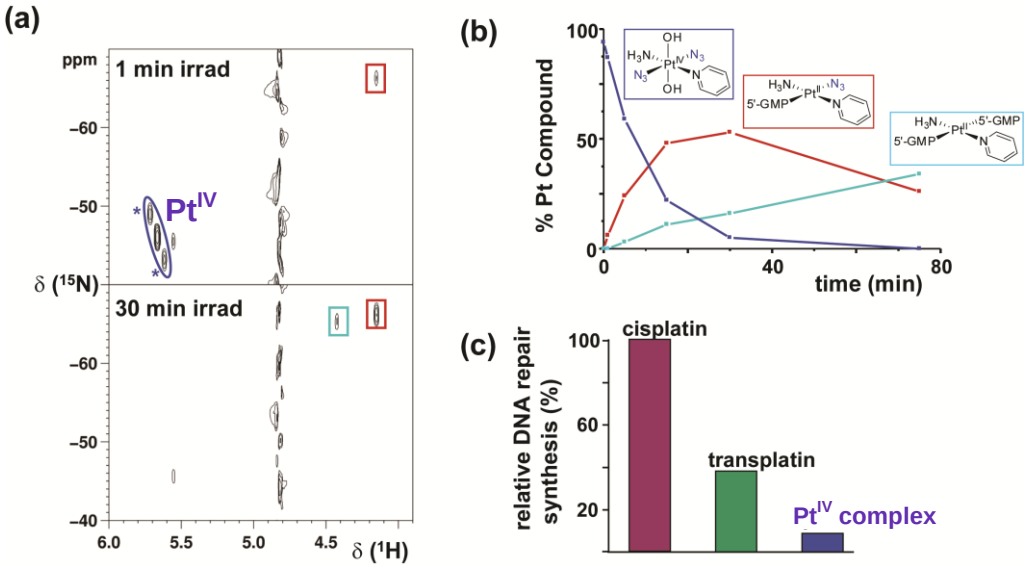


Figure 1-8. (a) [^{1}H,^{15}N] HSQC NMR spectra of ^{15}N-labeled ptohoactive Pt(IV) complex, *trans,trans,trans*-[Pt(N_3)$_2$(OH)$_2$(py)(NH_3)], and 2 mol equiv of 5'-GMP after 1 min and 30 min of irradiation; (b) time dependent decrease in concentration of *trans,trans,trans*-[Pt(N_3)$_2$(OH)$_2$(py)(NH_3)] and formation of [Pt^{II}(5'-GMP-N7)(N_3)(NH_3)(py)]$^{+}$ and *trans*-[Pt^{II}(5'-GMP-N7)$_2$(NH_3)(py)]$^{2+}$; (c) quantification of in vitro DNA repair synthesis using an extract from the repair-proficient HeLa cell line. Cisplatin was taken as 100%. Adapted from reference 32.

reactive Pt(II) species with liberation of two unstable azideradicals, which rapidly decompose in water to form N_2 gas (Figure 1-7). Aquation of the Pt(II) compound and subsequent binding to DNA (and possibly proteins) leads to its cytotoxic effect. Studies on calf thymus DNA platination have shown that the rates of photodecomposition closely correlate with the rates of irreversible DNA platination, indicating direct reaction of photolysis products with DNA at notably faster rates than cisplatin, for which the platination rate is determined by slow hydrolysis steps. Studies in the presence of 5'-GMP have shown that rapid photoreduction results in the formation of the mono-adduct [Pt^{II}(5'-GMP-N7)(N_3)(NH_3)(py)]$^{+}$ and bis-adduct *trans*-[Pt^{II}(5'-GMP-N7)$_2$(NH_3)(py)]$^{2+}$ (Figure 1-8). The observation suggests stepwise dissociation of the two azides rather than

concerted reductive elimination of both azides as a result of photoreaction of Pt(IV) complexes. Several biological assays such as different levels of p53 protein and caspase activity and distinctive crosslinks upon DNA platination pointed to an unusual mechanism of action which is clearly different from cisplatin. Moreover, DNA repair synthesis of the platinated lesion induced by photoplatination was considerably lower compared to cisplatin or transplatin (Figure 1-8).[35]

A novel platinum(IV) diazido complex *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ lately discovered in our group is stable in solution and can be photoactivated with a low-dose (5 J cm^{-2}) of visible light (420 nm) to give potent, cytotoxic effects in a number of cell lines. In the presence of 5'-GMP, the major photoproduct is *trans*-$[Pt(py)_2(5'\text{-}GMP)_2]^{2+}$ as confirmed by ^{1}H and ^{195}Pt NMR spectroscopy.[36]

Photoactivation promotes the rapid formation of such bis-guanine adducts whereas in thermal reactions involving *trans* diamine complexes such *trans* cross-links are formed very slowly if at all. The unusual nature of such *trans*-GG crosslinks formed on DNA after photactivation of *trans*-diazido Pt(IV) complexes may contribute to their high potency towards cancer cells.[35]

1.4.3 Sequence-Dependent Kinetics and Thermodynamics of Reactions between Platinum Complexes and DNA[37]

The reaction between DNA and platinum complexes (such as cisplatin) is often multistep, consisting of aquation, electrostatic preassociation, nucleophilic substitution of one aqua ligand by a DNA guanine, and crosslink formation. Quantification of monoadducts and diadducts after reaction between salmon testes DNA and $[PtCl_2(^{3}H\text{-}en)]$ as a function of incubation and postincubation time suggests that both the rate of formation of the monoadducts and the rate of their

closure to diadducts are sequence-dependent.[38] Herein, I summarise the recent studies on the sequence-dependent selectivity of nucleobases in DNA. A better understanding of how various forms of polynucleotides affect the stepwise interaction with metal-based complexes is essential for the design of drugs to target specific DNA sequences or structures.

1.4.3.1 Why Does Cisplatin Prefer Guanine over Adenine?

In early competition studies with nucleosides and nucleotides, guanine bases are often found to be the preferred DNA binding sites of cisplatin. Experimental kinetic[39] and theoretical[40,41] studies suggest that hydrogen bonding can partly account for the guanine over adenine preference of *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ and *cis*-$[PtCl(NH_3)_2(H_2O)]^{+}$ and, furthermore, the calculated free energies of activation show that *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ discriminates much more strongly between guanine and adenine than does *cis*-$[PtCl(NH_3)_2(H_2O)]^{+}$.[40]

1.4.3.2 How Is the Reactivity of an Individual Guanine Affected by the Sequence Context in Duplex DNA?

Experiments suggest that monoadduct formation proceeds under kinetic control rather than under thermodynamic control. To explain sequence-selectivity of *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ binding to duplex DNA, a kinetic model has been proposed based on HPLC studies.[42] According to this model, the neighbouring bases intervene (i) by affecting the molecular electrostatic potential at the site of the N7 lone-pair of guanine, and (ii) by exerting different steric hindrances for the platinum complex to approach the N7 atom of guanine. Hydrogen bonding from platinum ligands, initially considered as the most likely cause of the sequence selectivity, has been shown to intervene in a rather sequence-unspecific manner by

calculations that reveal the strongest hydrogen bond in the transition state forms between a water ligand of platinum and the 5'-phosphate group.[43] This may explain why the reactivity of complexes *cis*-$[PtL(NH_3)_2(H_2O)]^{n+}$ (n = 1, 2) increases in the order $Cl^- < NH_3 < OH^- < H_2O$ matches closely the increasing hydrogen-atom donating capacity.[44] In addition, kinetic studies of monoadducts with *cis*-$\{PtCl(NH_3)_2\}^+$ bound to a specific guanine (G*) in the centre of a DNA duplex give an impressive demonstration of how the aquation rate depends on the nature of the flanking bases.[37,45,46]

1.4.3.3 Why Are Interstrand Cross-Links the Minor DNA Products Formed by Cisplatin?

In addition to structure of the cisplatin–DNA interstrand crosslinks, how the sequence affects the kinetics of their formation is important for understanding the mechanisms of the drugs. Surprisingly fast closure of the aqua monoadducts formed at $d(GC)_2$ sequences of duplex oligonucleotides to interstrand crosslinks ($t_{1/2} \approx 1$ min at pH 4.5 and 20 °C),[45,47] suggests that the slow formation of interstrand crosslinks after DNA platination with *cis*-$[PtCl_2(NH_3)_2]$ is due to slow aquation of the chlorido monoadducts rather than a slow crosslinking reaction. Additionally, the 3'-cytosine has been shown to slow down the reactivity of the adjacent guanine in monoadducts formation (i) by lowering the negative potential at the N7 atom of the adjacent guanine, and (ii) by sterically hindering the access to the N7 of guanine with the 4-amino group of cytosine.[42]

1.4.4 Structure, Function and Targeting of G-quadruplexes

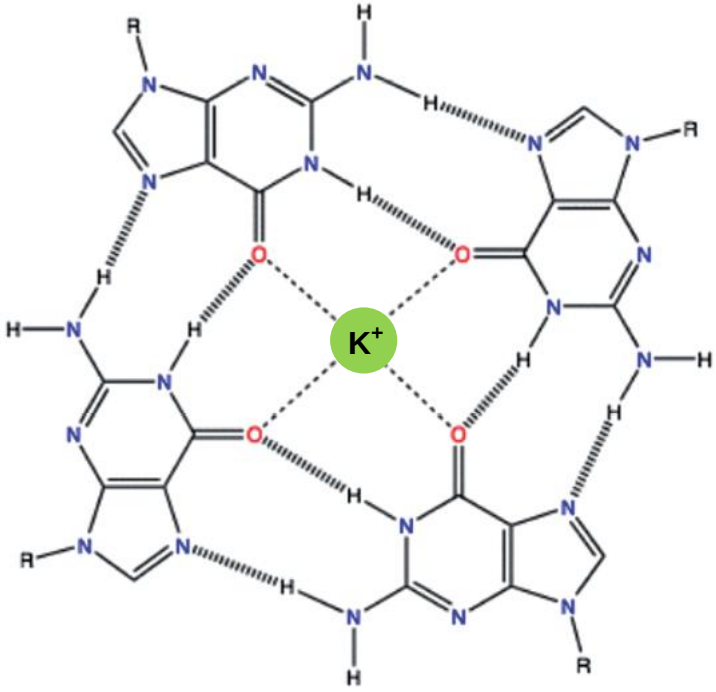


Figure 1-9. G-quartet with metal ion (e.g. K^+) coordination to GO6. Adapted from reference 48.

G-quadruplex nucleic acids can be formed by the spontaneous folding of DNA or RNA sequences that have repeat tracts of the kind AGnTm, into structures comprising stacked G-quartets held together by loops arising from the sequences occurring between the G-tracts, especially in the presence of a monovalent ion of the correct size, such as Na^+ or K^+ (Figure 1-9 and 1-10).[48] Quadruplex-forming sequences occur in particular at the ends of eukaryotic chromosomes, in promoter regions and/or in both 5′ and 3′ untranslated regions of a number of genes.[49]

Stabilization or disruption of quadruplex structures within their genomic environments can lead to a range of biological effects. Promoter quadruplex formation and stabilization can result in inhibition of the transcription of a targeted gene, and formation of telomeric quadruplexes can inhibit the end-

capping and catalytic functions of the telomerase enzyme from maintaining telomere integrity in cancer cells. G-quadruplex structures are, thus, the subject of considerable interest for the development of small-molecule ligands, and are also the targets of a wide variety of natural proteins.[49] More than 118 structures have been deposited in the protein data bank (January 2011).[50]

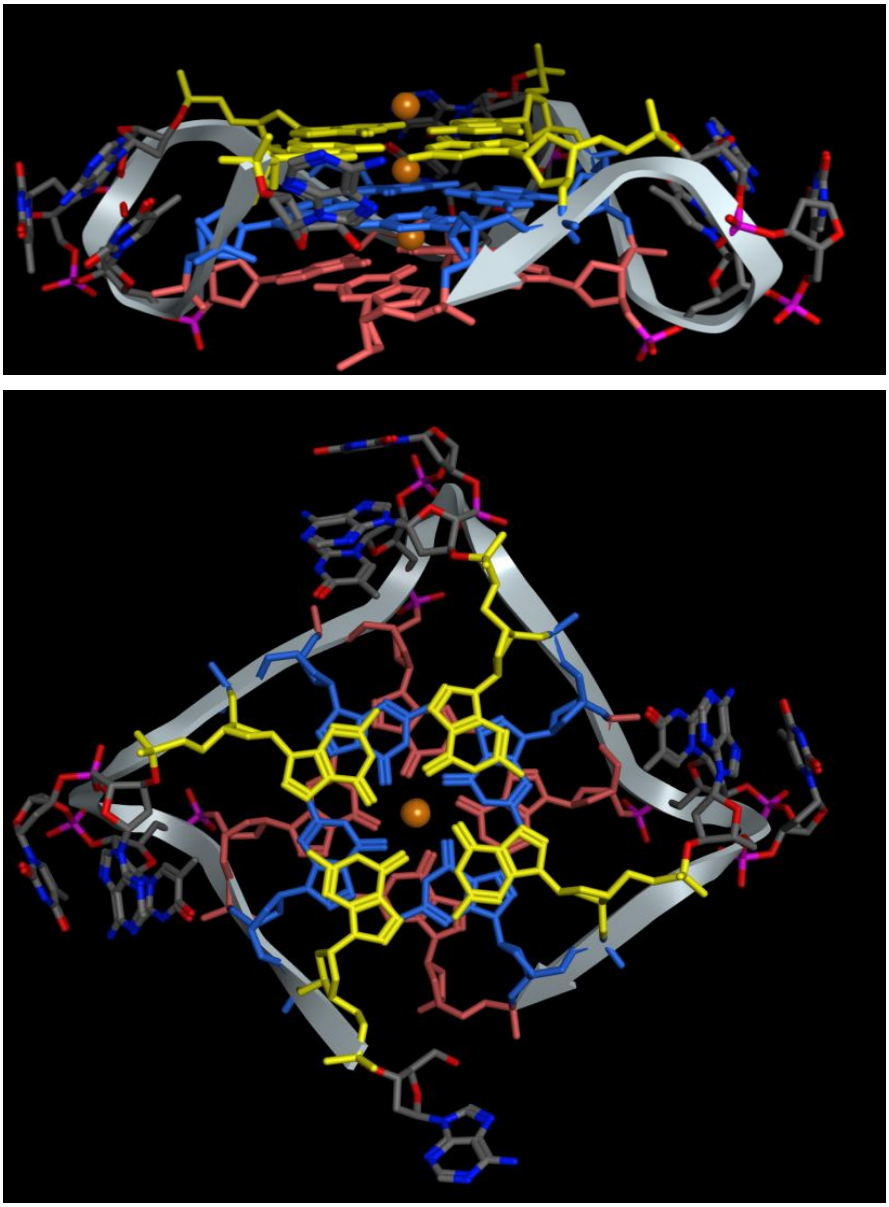
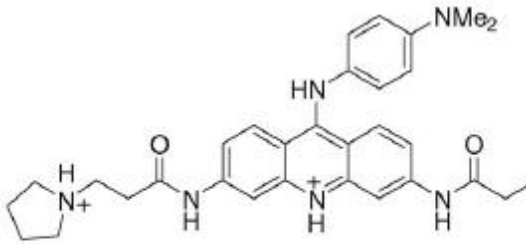
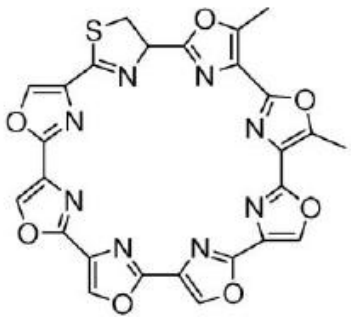


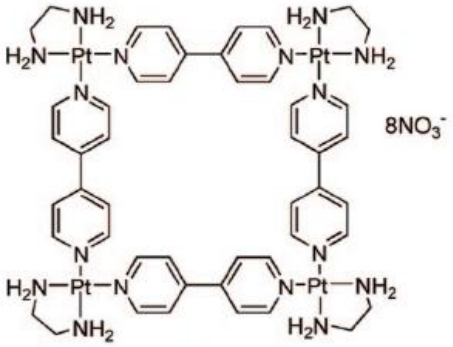
Figure 1-10. Visualizations of G-quadruplexes from PDB 1KF1. Ball = metal ion; white = DNA backbone; blue, yellow, and orange = Guanine. [51]



BARCO-19



Telomestatin



Platinum Supermolecular Square

Figure 1-11. Structures of three quadruplex-binding molecules.

Small molecules which bind to G-quadruplexes have been largely based on polycyclic planar aromatic compounds with at least one substituent terminating in a cationic group that enable them to intercalate effectively between planar G-quartets and generally bind more tightly to backbone phosphates of nucleic acids.[49] Some good quadruplex-binding molecules have been developed, such as those depicted in Figure 1-11. They include the trisubstituted acridine compound BRACO-19,[52] macrocycle telomestatin found in the bacterium *Streptomyces anulatus*,[53] and a platinum molecular square $[Pt(en)(4,4'\text{-dipyridyl})]_4$.[54] A particular challenge is for a given ligand to have high selectivity for any chosen one over other quadruplex structures, as well as to have low affinity for duplex DNA, which generally far outnumbers any G-quadruplexes present in the genome.[49,55] Structural information has revealed the presence of several distinctive topologies for telomeric quadruplexes, dependent on the nature of the cations present, and on length and sequence of the loops.[48,56]

1.5 Computational Approaches to Modeling Photoactivation and Platinum–DNA Cross-links

It is rare to see a natural product like taxol become a drug without any modification. The use of computational methods has shown advantage in the structure refinement and the prediction of pharmacological properties and of drug-receptor binding. The following section provides a collection of overviews of various theoretical methods used to study platinum anticancer complexes in this thesis. More advanced computational techniques are described in Chapter 2.

1.5.1 Time-dependent Density Functional Theory to Study Photochemistry and Molecular Spectroscopy

An extension of density functional theory (DFT) for excited states, time-dependent DFT (TDDFT), enables quantitative prediction of electronic absorption spectra for transition metal complexes. The results discussed in a recent review indicate that the agreement of TDDFT with experiment is not as good for transition metal complexes as has been observed for small molecules such as N_2, CO, CH_2O, and furan, where an accuracy of 0.1–0.2 eV has often been obtained. The best TDDFT methods for absorption spectra of transition metal complexes often appear to have errors of a few tenths of an eV. The applied functional and the geometry of the system are important for the accuracy of TDDFT.[62, 63]

For spectra in solution, considerable solvent shifts have been observed. Although various solvation models are available to include solvation effects in the TDDFT applications, the accuracy of solvation models in combination with TDDFT needs to be carefully evaluated.[62, 63] It is also important to note that the TDDFT module in current software suites such as the Amsterdam Density Functional (ADF)[57]

package is relatively new and not extensively tested for a wide range of applications. Therefore, one needs to build experience about aspects that may affect the accuracy of TDDFT results. In particular, it is advised to apply the zeroth-order regular approximation (ZORA)[58,59] relativistic corrections for molecules containing heavy nuclei and use an asymptotically correct exchange-correlation potential such as the statistical average of the orbital dependent potentials (SAOP)[60].

1.5.2 Previous Work on Photoactivated Pt(IV) Pyridyl Azido Acetato Complexes

The Pt(IV) complexes *trans, cis*-[Pt(bpy)$(OAc)_2(N_3)_2$] (**1**), *trans, cis*-[Pt(phen)$(OAc)_2(N_3)_2$] (**2**), *trans, trans, trans*-[Pt$(OAc)_2(N_3)_2(py)_2$] (**3**) were synthesised, and X-ray, NMR and UV-vis spectroscopy and TDDFT in the Gaussian 03 (G03) program[61] were used to study their structural and photochemical properties. The B3LYP functional[62,63] in conjunction with the LanL2DZ basis set[64] and effective core potential for the Pt atom and the 6-31G** basis set[65] for the other atoms, and the conductor-like polarisable continuum model method (CPCM)[66-68] with water as solvent were used to calculate the electronic structure and the excited states.

These complexes with aromatic amines undergo photoactivation with both UVA (365 nm) and visible green light (514 nm) and exhibit greater absorption at longer wavelengths (ε = 9756 M^{-1} cm^{-1} at 315 nm for **1**; ε = 796 M^{-1} cm^{-1} at 352 nm for **2**; ε = 16900 M^{-1} cm^{-1} at 307 nm for **3**, in aqueous solution) than previously reported Pt(IV) azide complexes. The calculated and experimental absorption spectra are displayed in Figure 1-12. Additional UV-Vis spectrum for complex **1** was

simulated using scalar relativistic TDDFT with the asymptotically correct XC potential derived from SAOP in the ADF program[57]. The computed spectra correlated with the experimental measurements, and the results show that TDDFT can be a useful tool to study the photochemistry of Pt(IV) complexes.

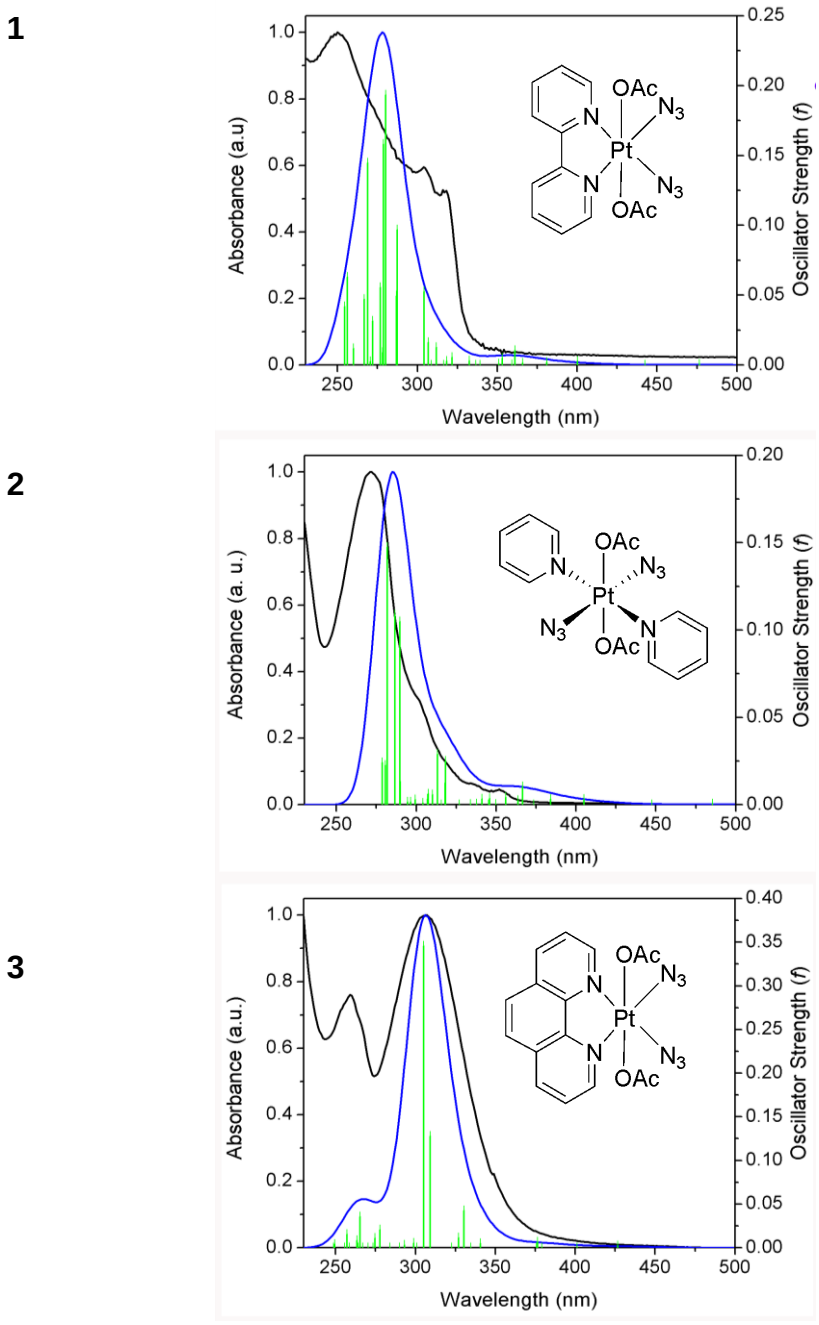


Figure 1-12. Calculated (blue line) and experimental (black line) absorption spectra of **1–3** in aqueous solution. The computed excited states are shown as vertical bars with heights equal to the extinction coefficients. Adapted from reference 69.

1.5.3 Molecular Modelling for Nucleic Acids

With significant progress in the past decade in the development of efficient methods for treating the long-range electrostatic interactions, empirical molecular mechanical force fields, and molecular dynamics (MD) simulation methods, current MD simulation is not only capable of reproducing experimental observations in general, but also can provide insight into the structure, energetics, and dynamics of nucleic acids. These force fields were mainly parameterised from experimental and quantum mechanical results on small molecular fragments that make up proteins and nucleic acids (see Chapter 2, Section 2.2 for more detailed description of force fields and parameterisation). Among available force fields, CHARMM and AMBER have been the most widely used. The performance of various force fields in the simulation of nucleic acids is provided in Section 2.2.4.

1.5.4 Molecular Modelling for Platinum–DNA Interactions

Herein, theoretical approaches to study platinum–DNA interactions are introduced. As computational models are often compared with experimental structures, cisplatin–DNA structures solved by NMR spectroscopy and X-ray crystallography as well as via quantum mechanics/molecular mechanics (QM/MM) methods are summarised.

1.5.4.1 Progress in the Application of Force Fields to Platinum–DNA Complexes

The diversity of structure makes inorganic molecules more difficult to model than their organic counterparts. A variety of force fields have been developed for the modelling of platinum–DNA complexes. The discrepancy between force field results and crystal structures of small molecules and oligonucleotide complexes as

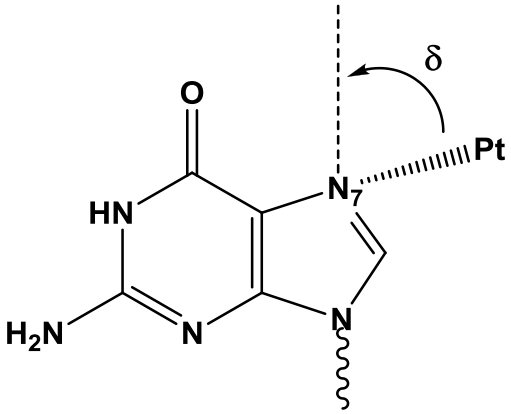


Figure 1-13. Deformation of Pt out of the plane of the guanine base. Adapted from reference 70.

well as geometries obtained by ab initio quantum mechanical calculations regarding the degree of bending of the Pt–N7 bond out of the plane of the coordinated purine base (δ, Figure 1-13) has been extensively discussed.[70] The parameters for the modelling of DNA interactions were developed by using torsional constraints or out-of-plane deformation terms to ensure the Pt atom remained within the plane of the coordinated purine. Ab initio calculations show that the assumption of planarity of nucleic acid bases leads to an overestimation of the platinum out-of-plane bending force constant.[71] They questioned whether the purine ring is stiff enough to withstand large out-of-plane bending of the Pt–N7 bond without puckering of the five-membered ring containing N7. The calculated force constants for puckering of the base and out-of-plane deviation of the Pt atom indicate that the purine bases are likely to undergo significant puckering. In the majority of platinum–DNA interactions, deviations from planarity are small and, therefore, the limitations in this aspect of the modelling are probably not significant, but modelling of highly strained systems may not be reliable. An explicit coupling between the geometry and the occupancy of d-orbitals can be

achieved by adding a ligand field energy term to the force field energy function, giving a more sophisticated approach to model transition metal complexes known as ligand field molecular mechanics (LFMM).[72]

1.5.4.2 Quantum Mechanics/Molecular Mechanics Approach

Hybrid quantum mechanics/molecular mechanics (QM/MM) methods have been widely applied to model local electronic events in large biomolecular systems. The basic idea is to describe the chemically active region such as platinated sites by electronic structure methods, as accurately as needed, while treating the rest by empirical force fields methods. This is the QM/MM approach based on a mixed Hamiltonian of the form:

$$\mathbf{H} = \mathbf{H}_{\mathrm{QM}} + \mathbf{H}_{\mathrm{MM}} + \mathbf{H}_{\mathrm{QM/MM}} \qquad (1\text{-}5)$$

$\mathbf{H}_{\mathrm{QM}}$ is the ab initio Hamiltonian, and $\mathbf{H}_{\mathrm{MM}}$ represents the classical Hamiltonian. In this way, large systems can be modelled in a computationally efficient manner. The main problems with QM/MM approach is deciding how the two regions should interact, as described by the interaction Hamiltonian $\mathbf{H}_{\mathrm{QM/MM.}}$ It is crucial to treat the QM/MM interface in an accurate and consistent way.

In general, there are three levels of $\mathbf{H}_{\mathrm{QM/MM}}$. The lowest level of the interaction is called *mechanical embedding* model, where only the bonded and steric energies of the two parts are considered, but the electronic interaction of the two regions is not included. This is not a widely applied model because the wave function of the QM region does not respond to structural changes in the MM region. In *electronic embedding* model, the atoms in the MM regions are allowed to polarize the QM region. Partial charges on the MM atoms can be incorporated into the QM Hamiltonian, and thus the QM atoms feel the electric potential caused by all the

MM atoms. As a result, the geometry of MM atoms can affect the QM region. In *polarisable embedding* model, the QM atoms are also allowed polarise the MM region. Thus, the electric field generated by the QM region influences the atomic charges and dipoles of the MM atoms. This model requires a polarisable force field and a double iterative procedure to compute the electric fields in both the QM and MM parts by a self-consistent protocol. This level of description demands a much larger computational cost, so it is not commonly used. Most QM/MM methods use the electronic embedding model.[73]

QM/MM data from Platts[74] used to compare with the LFMM results in Chapter 4 were computed by the ONIOM (own-n-layered integrated molecular orbital molecular mechanics) method, which, in general, uses mixing methods of different accuracy (e.g. two different inner QM layers and an outer MM layer). The structure of QM/MM simulation from Carloni[75] for the comparison of the resulting structure of the LFMD in Chapter 5 was obtained by *electronic embedding* model.

1.5.4.3 Molecular Dynamics Simulation on Platinum–DNA Complexes

Computer simulation methods enable us to model biological systems in laboratory-relevant conditions and predict their properties through generating representative configurations of these small replications in such a way that accurate values of structural and thermodynamic properties can be obtained with a feasible amount of computation.[76,77] Simulation techniques also enable the time-dependent behaviour of atomic and molecular systems to be determined, providing a detailed picture of the way in which a system changes from one conformation or configuration to another. Simulation techniques are widely used

in some experimental procedures, such as the determination of protein structures from X-ray crystallography.

A 5 ns MD simulation of the DNA duplex d(GCCG*G*ATCGC)• d(GCGATCCGGC) with a *cis*-$\{Pt(NH_3)_2\}^{2+}$ fragment crosslinking the two G* guanine bases was carried out using AMBER 98 force field[78] with reduced improper torsion angle terms that take into account purine puckering to parameterise the bending of the Pt-N7 bonds out of the guanine planes[71], and the results correlated with NMR data determined for the same adduct.[12] To resolve the differences between X-ray[79] and NMR[10] structures of d(CCTCTG*G*TCTCC)•d(GGAGACCAGAGG) containing a cisplatin coordinated to the two G* guanine bases, QM/MM simulations with the platinated unit treated by DFT and the rest of DNA by AMBER98 force field[78] were performed to study the same Pt–DNA adduct. The calculations were initiated with the X-ray structure of the platinated DNA, and during the 2-ns simulation the structure has shown closer resemblance to the solution structure obtained by NMR.[75]

1.5.4.4 Experimental and QM/MM Structures for Cisplatin–DNA Complexes

The structures of 1,2-intrastrand cisplatin–DNA complexes with the sequence d($C_1C_2T_3C_4T_5G_6$*G_7*$T_8C_9T_{10}C_{11}C_{12}$)• d($G_{13}G_{14}A_{15}G_{16}A_{17}C_{18}C_{19}A_{20}G_{21}A_{22}G_{23}G_{24}$) where the -$G_6$*$G_7$*- site has been modified by cisplatin are available in NMR (PDB 1A84)[10], X-ray at 2.6 Å resolution (PDB 1AIO)[79], and QM/MM calculation[75]. In AIO, however, T_3 was replaced by a 5-bromo-2'-deoxyuridine-5'-monophosphate to solve the crystal structure. The main structural features are summarised as follows. For NMR and X-ray structures, (i) DNA axis bend angles

are 78°, and 39°/55° (two independent molecules), respectively, (ii) minor groove width/depth are 9.4-12.5 Å/1.4 Å and 9.5-11.0 Å/3.0 Å, respectively, and (iii) the displacements of Pt atom from G6*/G7* are 0.8/0.8 Å and 1.3/0.8 Å, respectively. The QM/MM simulation started from the X-ray structure, and the resulting structure shows an increase of axis bend to an intermediate value between the X-ray and NMR structures. During the simulation, the structure of the platinated DNA rearranged significantly toward the NMR structure.

The structure of interstrand cisplatin–DNA complexes with the sequence $[d(C_1A_2T_3A_4G_5{*}C_6T_7A_8T_9G_{10})]_2$ cross-linked at the G*C:G*C site was determined by NMR (PDB 1DDP)[80] and that with the sequence $d(C_1C_2T_3C_4G_5{*}C_6T_7C_8T_9C_{10})\bullet d(G_{11}A_{12}G_{13}A_{14}G_{15}{*}C_{16}G_{17}A_{18}G1_9G_{20})$, where G* denotes the location of platinated nucleotides was solved by X-ray diffraction at 1.63 Å resolution (PDB 1A2E)[9]. The NMR structure exhibits a duplex bending of 14° toward the minor groove and a helix unwinding of 38°, while the X-ray structure shows a bend of 47° toward the minor groove and an unwinding of 70°. Note that the comparison of the structures needs to relate to the structural parameters used in Curve[81] program.

1.6 Advances in Current Status of *in Silico* Design of Photoactive Platinum Anticancer Complexes in the Thesis

Experimental and/or theoretical studies on photoactive platinum complexes and their mechanisms of action have been reported. However, no computational studies have been carried out to systematically tune the coordination environment to achieve desirable photochemical properties, and the elucidation of the cytotoxic

mechanism has been impeded by the difficulty of obtaining structures of platinum photoproducts with DNA/proteins via X-ray crystallography or NMR.

In this thesis, I first utilise TDDFT with the aim of designing platinum(IV) complexes with absorption bands shifted towards longer wavelengths. This is achieved via a broad exploration of *trans* influences, the nature of OR^- and (pseudo)halogen ligands, electron-withdrawing/donating/delocalizing substituents on the N-ligands, and intramolecular H-bonds. Secondly, LFMM parameters for Pt–guanine interactions are derived to model platinated DNA. Thirdly, the force field is used to develop protocols for LFMD simulations of known bifunctional DNA adducts involving both an intra- and an interstrand crosslink of the anticancer drug cisplatin. The fourth objective is to simulate a DNA lesion involving the photoproduct of the novel platinum(IV) complex *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ (**1**) and suggest possible mechanisms to explain the high potency of this photoactivated complex.

Chapter 2

Methods and Computational Details

This chapter provides an overview of the theoretical methods and techniques applied in this work. The first section provides an introduction to time-dependent density functional theory. Secondly, an overview of molecular mechanics, ligand field molecular mechanics, and molecular dynamics simulation is given. Thirdly, the ligand field molecular mechanics approach for modelling metal complexes is presented. Computational details are provided in the following section. Finally, the techniques used to analyse the results of molecular dynamics simulation are summarised.

2.1 Time-dependent Density Functional Theory

Time-dependent DFT (TDDFT) extends ground-state density functional theory (DFT) to the dynamical description of atomic and molecular electronic excitations and is an efficient computational tool for the study of molecular spectroscopy.[82,83] TDDFT can be viewed as an alternative formulation of time-dependent quantum mechanics, where the fundamental variable is no longer the many-body wave function but the one-body electron density. With this method, the time evolution of electronic systems can be simulated, which allows changes in molecular properties over time under the influence of applied time-dependent fields to be studied. Within TDDFT two main types can be distinguished: (i) If the external time-dependent potential is ‘strong’ (e.g. when matter interacts with intense laser fields), a full solution of the time-dependent Kohn-Sham equations has to be performed. (ii) For a ‘small’ external potential (e.g. for the calculation of photoabsorption spectra), the complete numerical solution of the time-dependent Kohn-Sham equations can be avoided by the use of linear response theory.

2.1.1 Basic Equations of TDDFT and Linear Response Theory

TDDFT linear response theory is one of the most widely used electronic structure method for calculating vertical electronic excitation energies.[84] The estimated accuracy of the best TDDFT methods for transition metal complexes, where errors are often of several tenths of an eV, is lower than for simple molecules like N_2, CO, and CH_2O, with errors of the order of 0.1–0.2 eV. According to the theorem of Runge and Gross[85], the time evolution of the electron density $\rho(\mathbf{r}_1,t)$ of a many-electron system at the initial time t=0 can be represented by a definite wavefunction $\Psi(\mathbf{x}_1,...,\mathbf{x}_N,t=0)$ ($\mathbf{x}_i$ are the electron coordinates). That is, given the initial state of the wavefunction, the density $\rho(\mathbf{r}_1,t)$ at all later time determines the potential $v_{ext}(\mathbf{r}_1,t)$ uniquely up to a purely time-dependent function. With the potential $v_{ext}(\mathbf{r}_1,t)$, the wavefunction $\Psi(\mathbf{x}_1,...,\mathbf{x}_N,t)$ of the evolving system can therefore become a functional of the time-dependent electron density, $\Psi(\mathbf{x}_1,...,\mathbf{x}_N,t)=\Psi([\rho],t)$. In other words, there is a one-to-one correspondence between $\rho(\mathbf{r}_1,t)$, $v_{ext}(\mathbf{r}_1,t)$, and $\Psi(\mathbf{x}_1,...,\mathbf{x}_N,t)$:

$$\rho(\mathbf{r}_1,t) \leftrightarrow v_{ext}(\mathbf{r}_1,t) \leftrightarrow \Psi(\mathbf{x}_1,...,\mathbf{x}_N,t) \tag{2-1}$$

In the Kohn-Sham (KS) approach of TDDFT,[86] the density ρ can be expressed in the orbital form:

$$\rho(\mathbf{r}_1,t) = \sum_i^N |\phi_i(\mathbf{r}_1,t)|^2 \tag{2-2}$$

by the sum of the densities of the KS orbitals ϕ_i, the first N solutions of the time-dependent one-electron KS equations (in atomic units)

$$\left\{-\frac{1}{2}\nabla^2 + v_s(\mathbf{r}_1,t)\right\}\phi_i(\mathbf{r}_1,t) = i\frac{\partial}{\partial t}\phi_i(\mathbf{r}_1,t) \tag{2-3}$$

The orbitals $\phi_i(\mathbf{r}_1,t)$ evolve under the influence of the KS potential v_s, which consists of the external potential v_{ext}, the standard Hartree potential v_{Coul} of the electrostatic electron repulsion and the exchange-correlation (xc) potential v_{xc}:

$$v_s(\mathbf{r}_1, t) = v_{ext}(\mathbf{r}_1, t) + v_{Coul}(\boldsymbol{r}_1, t) + v_{xc}(\boldsymbol{r}_1, t) \quad (2\text{-}4)$$

The time-dependent KS Equation (2-3) can be used to simulate the full time evolution of the density. However, if the external potential is small, the behaviour of the system can be described by first-order perturbation theory. The linear response approach considers a linear density response of the ground-state density to an external time-dependent perturbation $\delta v_{ext}(\mathbf{r}_1,t)$, for instance, an oscillating electric field $\delta v_{ext}(\mathbf{r}_1,t)$ switched on at time t_0:

$$\delta v_{ext}(\mathbf{r}_1, t) = \delta v_{ext}(\mathbf{r}_1)\cos(\omega t) \quad (2\text{-}5)$$

The time-dependent KS system obeys the time-dependent one-electron Schrödinger equation

$$\left\{-\frac{1}{2}\nabla^2 + v_s(\boldsymbol{r}_1) + \delta v_s(\mathbf{r}_1, t) - i\frac{\partial}{\partial t}\right\}[\phi_i(\mathbf{r}_1)e^{-i\varepsilon_i t} + \delta\phi_i(\mathbf{r}_1, t)] = 0 \quad (2\text{-}6)$$

where ε_i are orbital energies corresponding to KS molecular orbitals ϕ_i. Through Fourier transformation, the linear density response for the frequency-dependent density response $\delta\rho(\mathbf{r}_1,\omega)$ to the potential change $\delta v_s(\mathbf{r}_2,\omega)$ can be expressed as

$$\delta\rho(\mathbf{r}_1, \omega) = \int \chi^s(\mathbf{r}_1, \mathbf{r}_2, \omega)\delta v_s(\mathbf{r}_2, \omega)\, d\mathbf{r}_2 \quad (2\text{-}7)$$

and

$$\chi^s(\mathbf{r}_1, \mathbf{r}_2, \omega) = \sum_i^N \sum_{j\neq i}\left\{\frac{\phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)\phi_i(\mathbf{r}_2)}{\omega-(\varepsilon_j-\varepsilon_i)} - \frac{\phi_i(\mathbf{r}_1)\phi_j^*(\mathbf{r}_1)\phi_j(\mathbf{r}_2)\phi_i^*(\mathbf{r}_2)}{\omega+(\varepsilon_j-\varepsilon_i)}\right\} \quad (2\text{-}8)$$

ϕ_j and ϕ_i being the Kohn-Sham unoccupied and occupied molecular orbitals corresponding to the Kohn-Sham energies ε_j and ε_i, respectively. By defining δP_{ij} as

$$\delta P_{ij}(\omega) = \chi_{ij}^s \delta v_s^{ji}(\omega) \quad (2\text{-}9)$$

the density change is given as:

$$\delta\rho(\mathbf{r}_1,\omega) = \sum_{i,j} \delta P_{ij}(\omega)\phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1) \tag{2-10}$$

The χ^s in Equation (2-8) can be written as

$$\chi^s(\mathbf{r}_1,\mathbf{r}_2,\omega) = \sum_i^N \sum_a \phi_i(\mathbf{r}_1)\phi_a(\mathbf{r}_1)\phi_i(\mathbf{r}_2)\phi_a(\mathbf{r}_2)\left\{\frac{1}{\varepsilon_i-\varepsilon_a+\omega}+\frac{1}{\varepsilon_i-\varepsilon_a-\omega}\right\}$$

$$= \sum_i^N \sum_a \phi_i(\mathbf{r}_1)\phi_a(\mathbf{r}_1)\phi_i(\mathbf{r}_2)\phi_a(\mathbf{r}_2)\frac{2(\varepsilon_i-\varepsilon_a)}{(\varepsilon_i-\varepsilon_a)^2-\omega^2} \tag{2-11}$$

(from here we use *i* and *j* for the occupied orbitals, whereas *a* and *b* are used for the virtual ones). Then the elements $\chi_{ia}^s(\omega)$ in the density response function become:

$$\chi_{ia}^s(\omega) = \frac{(\varepsilon_i-\varepsilon_a)}{(\varepsilon_i-\varepsilon_a)^2-\omega^2} \tag{2-12}$$

Insertion of Equation (2-13) as well as the changes of external potential and of the potential of electron-electron interaction, Hartree and xc, into Equation (2-9) yields a more detailed equation:

$$\delta P_{ia}^{\sigma}(\omega) = \frac{(\varepsilon_{i\sigma}-\varepsilon_{a\sigma})}{(\varepsilon_{i\sigma}-\varepsilon_{a\sigma})^2-\omega^2}\left[\delta v_{ext}^{ia\sigma}(\omega) + 2\sum_{jb\tau} K_{ia\sigma,jb\tau}(\omega)\,\delta P_{jb}^{\tau}(\omega)\right] \tag{2-13}$$

(notice that we have now introduced the spin index σ and τ) where the coupling matrix consists of Hartree and xc parts:

$$K_{ia\sigma,jb\tau}(\omega) =$$

$$\int \phi_{i\sigma}^*(\mathbf{r}_1)\phi_{a\sigma}(\mathbf{r}_1)\left(\frac{1}{r_{12}} + f_{xc}^{\sigma\tau}(\mathbf{r}_1,\mathbf{r}_2,\omega)\right)\phi_{j\tau}(\mathbf{r}_2)\phi_{b\tau}^*(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{2-14}$$

with the xc kernel being the functional derivative of the xc potential $v_{xc\sigma}$ with respect to the spin-density ρ_τ

$$f_{xc}^{\sigma\tau}(\mathbf{r}_1,\mathbf{r}_2,t-t') = \frac{\delta v_{xc\sigma}(\mathrm{r}_1,t)}{\delta\rho_\tau(\mathrm{r}_2,t')} \tag{2-15}$$

Rearranging Equation (2-13) and separating its free terms leads to the following matrix equation for the density response

$$\sum_{\mathrm{jb}\tau}\left[M_{ia\sigma,jb\tau}(\omega)+\omega^2 S_{ia\sigma,jb\tau}\right]\delta P_{jb}^{\tau}(\omega)=\ \delta v_{ext}^{ia\sigma}(\omega) \tag{2-16}$$

where $M_{ia\sigma,jb\tau}(\omega)$ are the following matrix elements

$$M_{ia\sigma,jb\tau}(\omega)=\delta_{\sigma\tau}\delta_{ab}\delta_{ij}((\varepsilon_{j\tau}-\varepsilon_{b\tau})-2K_{ia\sigma,jb\tau}(\omega), \tag{2-17}$$

$$S_{ia\sigma,jb\tau}=\frac{(\delta_{\sigma\tau}\delta_{ab}\delta_{ij})}{\varepsilon_{j\tau}-\varepsilon_{b\tau}} \tag{2-18}$$

Excitation energies in response theory are characterized as the poles of the response functions. At the true excitation energies ω_k, a finite external perturbation $v_{\mathrm{ext}}(\mathbf{r}_1, \omega_k)$ leads to an infinite change in the density. For Equation (2-16) to hold, the matrix in the square brackets must to have a zero eigenvalue at ω_k. The following eigenvalue equations for the excitation energies ω_k can be derived:

$$\Omega\mathbf{F}_k=\omega_k^2\mathbf{F}_k, \tag{2-19}$$

where

$$\Omega=-\mathbf{S}^{-1/2}\mathbf{M}\mathbf{S}^{-1/2}, \tag{2-20}$$

so that its matrix elements are

$$\Omega_{ia\sigma,jb\tau}=\delta_{\sigma\tau}\delta_{ab}\delta_{ij}(\varepsilon_{j\tau}-\varepsilon_{b\tau})^2+2\sqrt{(\varepsilon_{a\sigma}-\varepsilon_{i\sigma})}K_{ia\sigma,jb\tau}\sqrt{(\varepsilon_{j\tau}-\varepsilon_{b\tau})} \tag{2-21}$$

The eigenvectors $\mathbf{F}_k$ can be used to calculate the corresponding oscillator strengths of the excitations.

2.1.2 The xc Potential

From the equations described in the previous section, it is evident that the KS orbital energies and the MO shapes are important for the xc potential and the

response properties such as excitation energies. Since the exact form of the xc potential $v_{xc\sigma}$ is not known, various approximations are in use. The local-density approximation (LDA) and the generally applied generalized gradient approximations (GGA) are the standard approximations in TDDFT. However, the LDA and GGA Kohn-Sham potentials have two main drawbacks. First, they are too shallow in the molecular region. Thus, the energies of unoccupied orbitals close to zero energy cannot shift up much and excitations to these levels tend to be underestimated in LDA or GGA potentials. Second, the KS potentials for LDA and GGA do not have correct Coulombic asymptotics $-1/r$. The much too fast exponential decay of these approximate potentials distorts the shape of diffuse unoccupied orbitals, and hence gives errors in intensities.

A new approximate xc potential, $\mathrm{v}_{\mathrm{xc}}^{\mathrm{SAOP}}$, constructed by the statistical average of (model) orbital potentials (SAOP)[60] brings a significant improvement in the calculated molecular static isotropic polarizabilities ($\mathrm{v}_{\mathrm{xc}}^{\mathrm{LDA}} < \mathrm{v}_{\mathrm{xc}}^{\mathrm{BP}} < \mathrm{v}_{\mathrm{xc}}^{\mathrm{SAOP}} < \mathrm{v}_{\mathrm{xc}}$).[87] The more sophisticated SAOP potential $\mathrm{v}_{\mathrm{xc}\sigma}^{\mathrm{SAOP}}$ employs a modified potential of van Leeuwen and Baerends, $\mathrm{v}_{\mathrm{xc}\sigma}^{\mathrm{LB}a}$, as one of its components,[60] and uses different model potentials, $\mathrm{v}_{\mathrm{xc}i\sigma}^{\mathrm{mod}}$, for the different occupied KS orbitals, $\phi_{i\sigma}$, and combines them into a single xc potential with the statistical average over the orbitals:

$$\mathrm{v}_{\mathrm{xc}}^{\mathrm{SAOP}}(\mathbf{r}_1) = \sum_{\mathrm{i}}^{\mathrm{N}} \mathrm{v}_{\mathrm{xc}i\sigma}^{\mathrm{mod}}(\mathbf{r}_1) \frac{|\phi_{i\sigma}(\mathbf{r}_1)|^2}{\rho_\sigma(\mathbf{r}_1)} \tag{2-22}$$

The orbital potentials $\mathrm{v}_{\mathrm{xc}i\sigma}^{\mathrm{mod}}$ are obtained with the following equation:

$$\mathrm{v}_{\mathrm{xc}i\sigma}^{\mathrm{mod}}(\mathbf{r}_1) = \mathrm{e}^{-2(\varepsilon_{N\sigma}-\varepsilon_{i\sigma})^2} \mathrm{v}_{xc\sigma}^{LBa}(\mathbf{r}_1) + \left[1 - \mathrm{e}^{-2(\varepsilon_{N\sigma}-\varepsilon_{i\sigma})^2}\right] \mathrm{v}_{xc\sigma}^{GLLB}(\mathbf{r}_1) \tag{2-23}$$

$\nu_{xc\sigma}^{LBa}$ has the proper Coulombic asymptotics (i.e. $-1/r$) in the outer region, and the potential $\nu_{xc\sigma}^{GLLB}$ of Gritsenko, van Leeuwen, van Lenthe, and Baerends (GLLB)[88] correctly describes the atomic shell structure in the inner regions. Therefore, $\nu_{xc\sigma}^{SAOP}$ provides a balanced approximation to the KS potential $\nu_{xc\sigma}$ in all regions.

2.2 Molecular Mechanics: Empirical Force Field Methods

Molecular mechanics calculate the energy of a system only as a function of the nuclear positions without considering the electronic motions, and is thus often used to perform calculations on systems of biological interest that typically involve macromolecules in their condensed phase environment. In addition, to study the inherent dynamical nature of biomolecules and the mobility of their environments, energy calculations of a significant number of conformations are required. Molecular mechanics is based on empirical energy functions with relatively simple terms to describe the underlying physical interactions that direct the structure and dynamic properties of molecules and can therefore fulfill the demands required by computational studies of biological systems. Molecular mechanics, in some cases, can provide properties that are as accurate as even the highest-level quantum mechanical calculations, at a fraction of the computational cost. However, to model molecular properties that depend upon the electronic distribution, careful parametrisation is essential.

2.2.1 A Simple Molecular Mechanics Force Field

Most force fields designed to model the total potential energy can be interpreted by a relatively simple four-component picture of the intra- and inter- molecular forces within the system.[89] The intramolecular terms are associated with

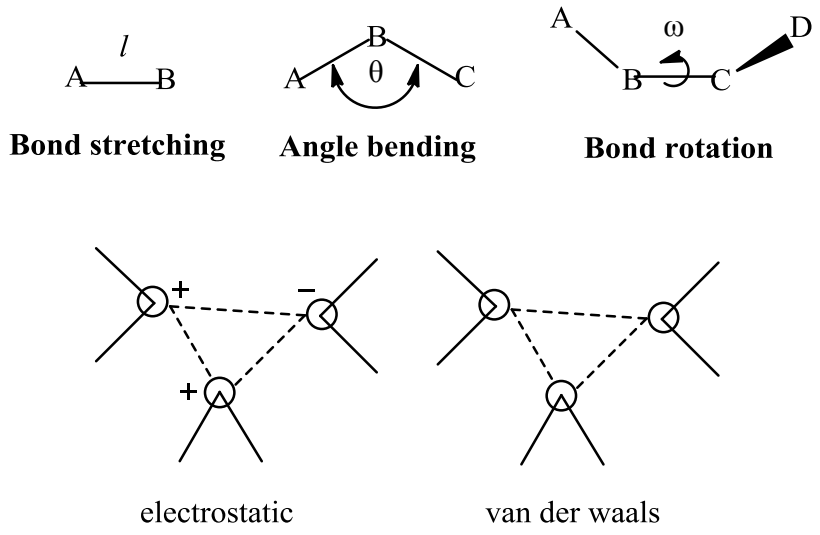


Figure 2-1. Illustration of the four principle contributions to a molecular mechanics force field: bond stretching, angle bending and torsional terms and non-bonded interactions.

covalently connected atoms, and the intermolecular terms represent the noncovalent or non-bonded interactions between atoms. The total potential energy of a chemical system, $E(\mathbf{r}^N)$, can be described as a function of the positions (**r**) of N particles in the following equations:

$$E(\mathbf{r}^{\mathrm{N}}) = \sum_{\mathrm{bonds}} \frac{k_{li}}{2}\left(l_i - l_{i,0}\right)^2 + \sum_{\mathrm{angles}} \frac{k_{\theta i}}{2}\left(\theta_i - \theta_{i,0}\right)^2 + \sum_{\mathrm{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma)) + \sum_{\mathrm{i=1}}^{N} \sum_{\mathrm{j=i+1}}^{N} \left(4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right) \quad (2\text{-}24)$$

The various contributions are illustrated in Figure 2-1. The first term in Equation (2-24) models a bond between two atom types A and B using a harmonic potential that gives the increase in energy as the bond length l_i deviates from the reference value $l_{i,0}$. The second term is a summation over every valence angle formed

between three atoms A–B–C in which A and C are both bonded to B in the molecule, modelled here by a simple harmonic approximation. The third term is a torsional potential that models how the energy changes by rotation around a bond. The fourth contribution is the non-bonded term that describes interactions between all pairs of atoms (*i* and *j*) that are in different molecules or that are in the same molecule but separated by at least three bonds. A basic force field expression usually models the non-bonded term using a Coulomb potential for electrostatic interactions and a Lennard-Jones potential for van der Waals interactions.

More sophisticated force fields may have a number of additional components, but they invariably contain these four energy terms. An advantage of this representation is that it is easier to understand how changes in the force field parameters affect its performance. This is helpful for the parameterisation as the various terms can be attributed to changes in specific internal coordinates such as bond lengths, bond angles, the rotation of bonds or movements of atoms relative to one another.

2.2.2 Force Field Parameterisation

Parameterisation of a force field is by no means a trivial task. A force field can contain a large collection of functional forms and associated constants, even for calculations on only a small molecular system. Creating a new force field entirely from scratch takes tremendous effort. It can be a complicated and time-consuming process even to model a new class of molecules by adding a few parameters to an existing force field. It is noteworthy that the performance of a force field is often particularly sensitive to a few of the parameters, and thus it is wise to spend more

time optimising these parameters rather than the values which do not greatly affect the results.

The first step in the parameterisation process is to select the data (e.g. conformations and vibrational frequencies) that one wishes the force field to reproduce. In the cases that experimental data are not available or difficult to obtain, quantum mechanics calculations are often used to provide the data for the parameterisation of molecular mechanics force fields. This significantly extends the range of chemical systems that can be modelled by the force-field approach.[89]

2.2.3 Force Field Models for Liquid Water

Proper condensed phase simulations require that the non-bonding interactions between different solvents and solutes of the system under study be properly balanced. The first step in obtaining this balance is to model the solvent–solvent interactions. Despite its small size, water is one of the most challenging systems to model accurately. Many different water models have been proposed. As there may be a very large number of water molecules present, most of the force fields used to simulate liquid water use effective pairwise potentials with no explicit three-body terms or polarisation effects. The TIP3P[90] model, one of the most common water models in biomolecular simulations, uses pairwise Coulombic and Lennard-Jones expressions for the interaction between molecules and a total of three interaction sites for the electrostatic interactions with the partial positive charges on the hydrogen atoms exactly balanced by an appropriate negative charge located on the oxygen atom in a rigid geometry. A Lennard-Jones function with simply a single interaction point per molecule centred on the oxygen atom is employed to model the van der Waals interaction between two water molecules. No van der Waals

interactions involving the hydrogen atoms are computed. The simple water models are usually parameterised by calculating various thermodynamic and structural properties (e.g. the density, dipole moments, radial distribution function, heat capacity, diffusion coefficient, and dielectric constant) using molecular dynamics or Monte Carlo simulations and then modifying the parameters until the desired level of agreement between experiment and theory is achieved.[89] Depending on the properties of water to reproduce in a system, an appropriate model needs to be selected with care.

2.2.4 Force Fields for Nucleic Acids

Both widely applied force fields, AMBER94[91] and CHARMM22[92], can reproduce many static properties of nucleic acids and yield stable unconstraint MD structures in physiological conditions. CHARMM27[93,94] has corrected the overstabilisation of A forms relative to B forms of DNA in CHARMM22. AMBER99[78] has improved sugar pucker phases, endocyclic torsion chi values, and overall helical twist, but fails to show B-DNA to A-DNA transitions under conditions expected to stabilise A-DNA. In contrast, this transformation is observed with AMBER94. Although perform well in simulations in the 10-ns range, AMBER94 and AMBER99 have been reported to overpopulate the α/γ backbone in the *gauche+* , *trans* geometry which, may introduce severe distortions in DNA in long simulation. parmbsc0[95], a new AMBER force field, has improved the description of α/γ conformers and can be used to study nucleic acids in a greater time scale. A cautionary note regarding the relative performance of the force fields is that modifications to a given force field do not necessarily improve all properties. It is also noteworthy that the comparison to experimental structure values tends to be biased by poorly refined NMR structures or crystal packing artifacts.

In solution, the DNA duplex is a highly charged polyanion stabilised by a range of cations. Counterions and solvation conditions play important roles in conformational properties and function of DNA, and thus a proper description of the ions around DNA is essential for accurate modeling of the dynamical structure of DNA in MD simulations.[96] The results for MD simulations of DNA dodecamer together with water molecules and Na^+ ions show that the internal structural parameters of DNA are well converged around 5 ns, yet the relaxation times of ion motions are much longer.[97]

2.2.5 Force Fields for Metal Coordination Complexes

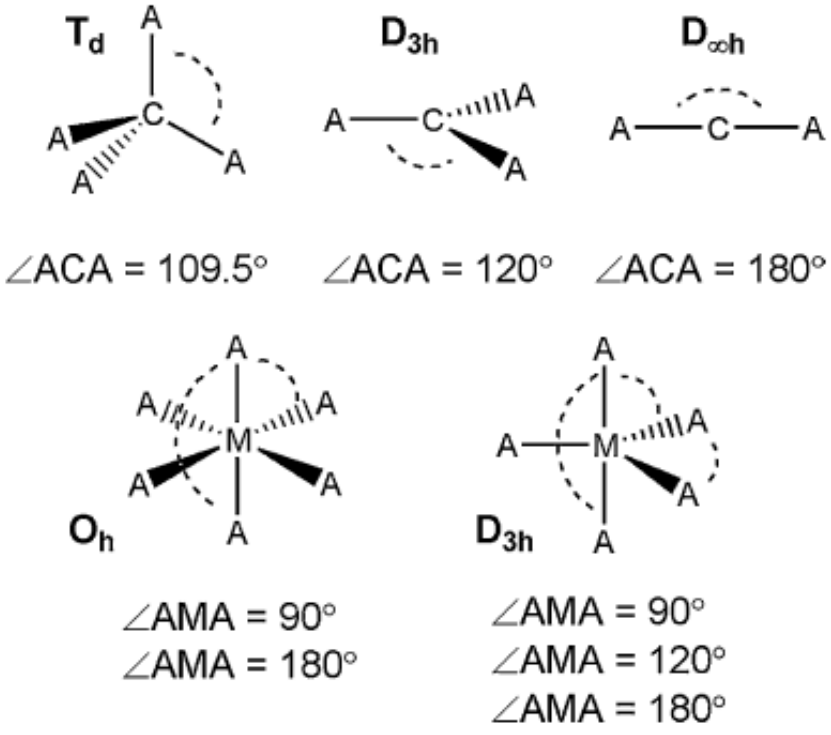


Figure 2-2. Valence angles at central atom for various regular geometries. Reproduced from reference 72 with permission.

The diversity of geometries and the presence of highly delocalised bonds make inorganic molecules more difficult to model than their organic counterparts.

Common structures of metal complexes include tetrahedral and square planar for four coordination, square pyramidal and trigonal bipyramidal for five coordination, and octahedral and trigonal prismatic for six coordination. Carbon coordination chemistry is straightforward. As the three common geometries — tetrahedral, trigonal planar and linear — are each associated with a single angle — 109.5◦, 120◦, and 180◦, respectively, the reference value, $\theta_{i,0}$, in Equation (2-24) for the bond angle term is clear. In contrast, for common coordination geometries like octahedral and trigonal bipyramidal, there are multiple reference angles for the same A–M–A triad (Figure 2-2). The increased number of ligands combined with the multitude of possible geometries significantly increases the problem of assigning suitable functional forms for each of the energy terms. Hence, the first hurdle encountered in molecular mechanics for transition metal systems is how to describe the angular geometry at the metal centre when using a force field expression like Equation (2-24). The situation can be much more challenging for irregular four-, five- and six-coordinate species particularly when valence electronic effects such as the Jahn-Teller effect play a significant role in determining the molecular structure.[98] A Urey-Bradley[92] treatment of the bonding about the metal can often give quite reasonable geometries; here, there are no angle-bending terms at the metal but terms due to pairs of atoms bonded to the metal. Alternatively, ligand field molecular mechanics (LFMM)[99] — a more sophisticated approach that includes an explicit treatment of d-electron effects — has been developed and successfully applied to the modelling of various transition metal complexes.

2.3 Ligand Field Molecular Mechanics

Ligand field molecular mechanics (LFMM)[72,98-101] is designed to model the electronic effects of the valence d and s electrons in transition metal compounds. These effects can lead to severe distortions away from the ideal geometries predicted by simple VSEPR theory. The stereochemical activity of d electrons manifests in a range of structural distortions of ionic coordination complexes typified by the Jahn–Teller elongations of six-coordinate d^9 Cu^{2+} species. Modelling these effects requires an additional term in the potential energy which describes the ligand field stabilisation energy (LFSE). The LFSE is explicitly incorporated into the LFMM method which has been applied to a range of complexes of Cu^{2+}, Ni^{2+}, Co^{3+}, Ru^{2+}, and Pt^{2+}. A single set of LFMM parameters for a given metal–ligand interaction is able to model different coordination numbers and geometries, spin states and bond lengths. [72,98,100,102]

2.3.1 The Angular Overlap Model

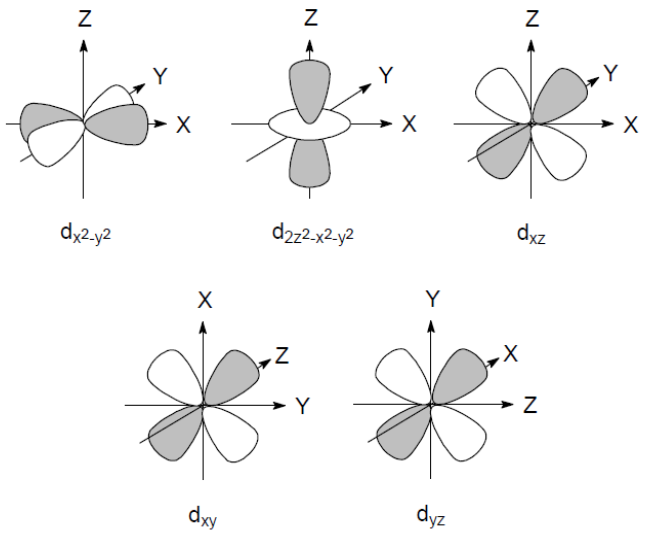


Figure 2-3. Schematic representation of the five d orbitals. Note change in axis frame directions for d_{xy} and d_{yz}.

The shapes of the d-orbitals are central to the qualitative and quantitative aspects of ligand field theory (LFT) (Figure 2-3). Certain magnetic and spectroscopic

features arise from the electrons in the d-shell which play a crucial role in the chemistry of the transition elements.

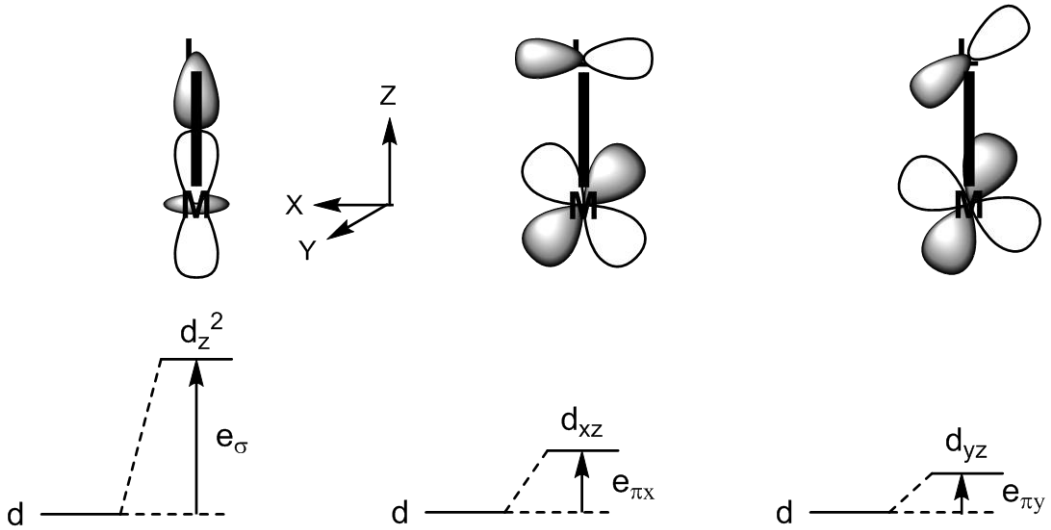


Figure 2-4. Definition of AOM parameters in terms of local M–L bonding. Reproduced from reference 72 with permission.

Although valence bond theory and crystal field theory (CFT) adequately account for a large number of features, they have serious limitations when describing the bonding in coordination compounds and rationalizing and predicting their properties.[103] The concept now most widely used to study the general chemistry of the transition metals is a simple molecular orbital (MO) approach, often referred to as Ligand Field Theory (LFT). The notions inherent in the simple MO approach form the basis of a potent method of describing metal–ligand bonding in terms of local σ and π bonding interactions while retaining the empiricism of the ligand field approach. It is known as the angular overlap model (AOM).[104] The AOM defines the total ligand field potential as a sum of individual, local M–L σ and π bonding contributions as illustrated in Figure 2-4, rather than in terms of the symmetry of the complex as a whole. The AOM parameters e_σ, $e_{\pi x}$, and $e_{\pi y}$ may

be derived from spectroscopic measurements or fitted to the results of quantum mechanical calculations .[100]

2.3.2 The LFMM Method

LFMM uses a generalized ligand field calculation of the d-orbital energies to estimate the ligand field stabilisation energy (LFSE). The LFSE depends on the complex, the d configuration and the spin state and can have a marked effect on a range of structural and thermodynamic properties of metal complexes. Since MM is inherently bond centered, the most convenient form of ligand field model is the angular overlap model (AOM) outlined above.[104] Each M-L interaction is parameterized in terms of local bonding modes of σ or π symmetry, with each assigned its own parameter (Figure 2-4).[100] The full ligand field potential is then constructed by summing the contributions from each bonding mode over all ligands and diagonalizing the resulting matrix. The LFSE and its gradients are then merged with a more-or-less conventional MM scheme. The total potential energy, $E(\mathbf{r}^N)$, can be modeled by simple contributions from bond stretching, E_{str}, angle bending, E_{bend}, torsional twisting, E_{tor}, and non–bonding interactions, E_{nb}, where the latter may include both van der Waals and electrostatic terms. The core concept of the ligand field molecular mechanics (LFMM) method is the addition to the conventional potential energy expression of a new term specifically designed to describe the effects of the ligand field stabilization energy (LFSE):

$$E(\mathbf{r}^N) = \Sigma\, E_{\text{str}} + \Sigma\, E_{\text{bend}} + \Sigma\, E_{\text{tor}} + \Sigma\, E_{\text{nb}} + \text{LFSE} \qquad (2\text{-}25)$$

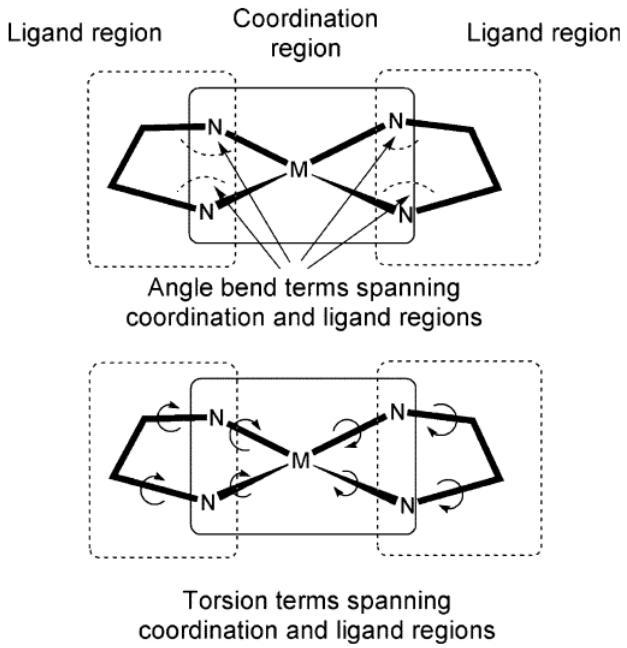


Figure 2-5. Schematic reprecsentation of division into coordination and ligand regions and force field terms which span the two. Reproduced from reference 72 with permission.

In other words, LFMM merges conventional MM for the 'organic' parts of a transition metal complex with an AOM treatment of the LFSE for the metal centre. Each molecule is divided into two overlapping regions as illustrated in Figure 2-5 for a $[M(en)_2]^{n+}$ complex (en = ethylenediamine). The coordination region contains the metal and its bonded donor atoms (e.g. the MN_4 unit) and the ligand region includes everything except the metal atom (e.g. the ethylenediamine ligands). The LFMM model focuses on the coordination region and is used to treat the LFSE, M–L bond stretching (via a Morse function), and L–M–L angle bending (via an explicit L-L repulsive term). The program, 'd orbital molecular mechanics in the Molecular Operating Environment' (DommiMOE)[101,105], has been designed to take advantage of existing force fields implemented in MOE, e.g., MMFF[106-110], AMBER[111] and CHARMM[112], which treat all the interactions in the ligand region.

2.4 Molecular Dynamics Simulation Methods

Molecular dynamics (MD) simulation, which provides detailed microscopic modelling on the atomic scale by numerically solving Newton's equations of motion, is an excellent approximation for a wide range of systems and is thus a powerful tool in chemistry, physics, biology, and materials science. This technique enables the natural time evolution of the system which allows the structural, dynamic, and thermodynamic properties of substances to be predicted directly from the underlying interactions between the molecules, providing every detail of how a many-particle system changes from one conformation to another.[113, 114, 89]

2.4.1 Basic Statistical Mechanics of Molecular Dynamics

Computer simulation generates information at the microscopic level (atomic and molecular positions, velocities etc.) and the conversion of this very detailed information into macroscopic properties (pressure, internal energy etc.) is the province of statistical mechanics. Herein, the aspects of the subject which are within the scope of this thesis are summarised.[89,113]

A Hamiltonian, *H*, in a classical system is a function of both coordinates **r** and momenta **p**. The Hamiltonian for the system can be expressed as the sum of kinetic energy, *K*(**p**), and the potential energy, *U*(**r**),

$$H = H(\mathbf{r}, \mathbf{p}) = K(\mathbf{p}) + U(\mathbf{r}) = \sum_i \frac{p_i}{2m_i} + U(\mathbf{r}) \quad (2\text{-}26)$$

where p_i is the momentum of particle *i*, and m_i the mass of particle *i*. A microscopic state of the system can therefore be characterised by the set of values {**r, p**}. For a system containing *N* atoms, 3*N* positions and 3*N* momenta are

required to define the state of the system. The 6*N*-dimensional space defined by by both coordinates **r** and momenta **p** is known as the *phase space* of the system.

For a system with fixed number of particles, volume and temperature, the probability density is given by the Boltzmann distribution function:

$$\rho(\mathbf{r},\mathbf{p}) = \frac{\exp\left(-H(\mathbf{r},\mathbf{p})/k_B T\right)}{Z} \tag{2-27}$$

where the canonical partition function, *Z*, is an integral over all phase space of the Boltzmann factors $\exp\left(-H(\mathbf{r},\mathbf{p})/k_B T\right)$, and k_B is the Boltzmann constant. Partition function allows us to compute thermal averages of any dynamic variable A(**r, p**) of interest (positions, the total energy, the kinetic energy etc.) and, therefore, is the key feature in statistical mechanics. These averages over every possible state of the system:

$$\langle A(\mathbf{r},\mathbf{p})\rangle_z = \int_V d\mathbf{r} \int_{-\infty}^{\infty} d\mathbf{P}\rho(\mathbf{r},\mathbf{p})A(\mathbf{r},\mathbf{p}) \tag{2-28}$$

are called *ensemble averages*. Alternatively, to calculate the average values of properties, averaging over the whole phase space all at once can be replaced by following the motion of a single point through phase space. That is, in this approach, the motion of a single point through phase space is followed as a function of time, and the averages, known as *time averages*, simply involve computing those points during the time evolution:

$$\langle A(\mathbf{r},\mathbf{p})\rangle_\tau = \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau A(\mathbf{r}(t),\mathbf{p}(t))\, dt \tag{2-29}$$

where τ is the duration of the simulation. If τ is sufficiently long, the value of the integral in Equation (2-29) approaches the 'true' average value of the property. Compared to the ensemble average, the time average is easier to obtain.

The *ergodic hypothesis*, one of the fundamental axioms of statistical mechanics, states that for an infinitely long trajectory the ensemble average is equal to the time average:

$$\lim_{\tau\to\infty}\langle \mathrm{A}(\mathbf{r},\mathbf{p})\rangle_{\tau} = \langle \mathrm{A}(\mathbf{r},\mathbf{p})\rangle_{Z} \tag{2-30}$$

In other words, the ergodic hypothesis assumes that when simulated time becomes long enough, eventually all of phase space will be explored and thus quantities by time averaging and by ensemble averaging become equal. The system has to reach equilibrium for this hypothesis to sustain.

2.4.2 Newtonian Molecular Dynamics[114]

In molecular dynamics, configurations of the system connected in time are generated by a numerical integration of Newton's laws of motion. As a result, a trajectory that specifies how the positions and velocities of the particles in the system evolve as a function of time is recorded. Newton's equation of motion states that

$$F_i = m_i a_i = m_i \frac{d^2 r_i}{dt^2} \tag{2-31}$$

where F_i is the force on particle i, m_i is the mass of particle i, a_i is its acceleration, and r_i is the coordinate.

In order to generate the trajectory, a numerical procedure for integrating the differential equation to solve Newton's equation of motion is required. A standard approach to solve ordinary differential equations like Newton's equation of motion is a finite difference method. The essential concept is that the integration can be broken down into many small steps with a fixed time interval Δt. Given the molecular coordinates and velocities at time t, the molecular coordinates and

velocities at a later time $t+\Delta t$ can be calculated to a sufficient degree of accuracy. The equations are solved step by step. The choice of time step Δt will depend on the characteristics of the simulated molecules, and Δt will be significantly (typically an order of magnitude) smaller than the typical time of the fastest motion studied in order to model such a motion with sufficient accuracy.

For algorithms using finite difference methods, the positions can be approximated as a Taylor expansion:

$$\mathbf{r}(t\ +\ \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 + \frac{1}{6}\mathbf{b}(t)\Delta t^3 + \frac{1}{24}\mathbf{c}(t)\Delta t^4 + \cdots \qquad (2\text{-}32)$$

where $\mathbf{v}(t)$ is the velocity (the first derivative of the positions with respect to time), and $\mathbf{a}(t)$ is the acceleration (the second derivative), $\mathbf{b}(t)$ is the third derivative, and so on. With $\mathbf{r}_n$ and $\mathbf{r}_{n+1}$ to indicate the position at step n (at time t) and at the next step, $n+1$ (at time $t+\Delta t$), respectively, and recalling Equation (2-31), Equation (2-32) can be written as:

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{v}_n\Delta t + \frac{1}{2}\left(\frac{\mathbf{F}_n}{\mathbf{m}}\right)\Delta t^2 + O(\Delta t^3) \qquad (2\text{-}33)$$

2.4.2.1 Verlet Integrator

The Verlet algorithm[115] is one of the simplest method for integrating the equations of motion in a molecular dynamics simulation. The Verlet integrator employs two Taylor expansions, a forward expansion $(t+\Delta t)$ and a backward expansion $(t-\Delta t)$:

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{v}_n\Delta t + \frac{1}{2}\left(\frac{\mathbf{F}_n}{\mathbf{m}}\right)\Delta t^2 + O(\Delta t^3) \qquad (2\text{-}34a)$$

$$\mathbf{r}_{n-1} = \mathbf{r}_n - \mathbf{v}_n\Delta t + \frac{1}{2}\left(\frac{\mathbf{F}_n}{\mathbf{m}}\right)\Delta t^2 - O(\Delta t^3) \qquad (2\text{-}34b)$$

Summing these two expansions, an algorithm for propagating the position with an error of order $O(\Delta t^4)$ is obtained:

$$\mathbf{r}_{n+1} = 2\mathbf{r}_n - \mathbf{r}_{n-1} + \left(\frac{\mathbf{F}_n}{\mathbf{m}}\right)\Delta t^2 + O(\Delta t^4) \quad (2\text{-}35)$$

Subtracting Equation (2-34b) from Equation (2-34a) gives a complementary algorithm for propagating the velocities accurate to order $O(\Delta t^2)$:

$$\mathbf{v}_n = \frac{\mathbf{r}_{n+1} - \mathbf{r}_{n-1}}{2\Delta t} + O(\Delta t^2) \quad (2\text{-}36)$$

The advantages of the Verlet algorithm are (i) implementation is straightforward and the storage requirements are modest, comprising two sets of positions ($\mathbf{r}_n$ and $\mathbf{r}_{n+1}$) and the accelerations $\mathbf{a}(t)$, (ii) fast since only a single force evaluation per integration cycle is required, and (iii) time reversibility. The disadvantages of the Verlet algorithm are (i) the positions $\mathbf{r}_{n+1}$ are obtained by adding a small term, $\mathbf{a}(t)\Delta t^2$, to the difference of two much larger terms, $2\mathbf{r}_n$ and $\mathbf{r}_{n-1}$ and this may lead to large numerical errors, (ii) $\mathbf{v}_n$ can be computed only if $\mathbf{r}_{n+1}$ is already known and the velocity propagation is only accurate to order $O(\Delta t^2)$ (an accurate estimate of the velocity is required to compute the kinetic energy and the instantaneous temperature), and (iii) it is not *self-starting* (the new positions, $\mathbf{r}_{n+1}$, are computed from the positions at the current, $\mathbf{r}_n$, and the previous, $\mathbf{r}_{n-1}$, time steps), and a lower order Taylor series expansion is often used to initiate the propagation.

2.4.2.2 Velocity Verlet Integrator

The *velocity Verlet* algorithm[116] has been developed to improve the velocity evaluation in the basic Verlet integrator. This algorithm gives positions, velocities, and accelerations all at the same time *t* and minimizes round-off errors by the following form:

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{v}_n\Delta t + \frac{1}{2}\left(\frac{\mathbf{F}_n}{\mathbf{m}}\right)\Delta t^2 \quad (2\text{-}37a)$$

$$\mathbf{V}_{n+1} = \mathbf{V}_n + \frac{1}{2}\left[\frac{\mathbf{F}_n}{\mathbf{m}} + \frac{\mathbf{F}_{n+1}}{\mathbf{m}}\right]\Delta t \qquad (2\text{-}37b)$$

The advantages of the velocity Verlet algorithm are (i) it is simple and numerically stable, and (ii) it is a more accurate equation for the velocities and consequently the kinetic energy compared to the Verlet algorithm. The main disadvantage of this method is the slightly higher computational cost compared to the simpler Verlet or leap-frog algorithms.

2.4.3 Setting up and Running a Molecular Dynamics Simulation

To start a molecular dynamics simulation, we first need to decide which energy model is to be used to describe the interactions within the system and specify the conditions of the run (e.g., number of particles, initial temperature). Secondly, we need to initiate the integration by selecting initial values of the system (typically a set of initial coordinates {r(0)} and a set of initial velocities {v(0)}). Following the assignment of an initial configuration is an *equilibration phase*, during which the system evolves from the starting configuration. Thermodynamic and structural properties must be monitored carefully during this stage until stability is achieved. At the end of the equilibration proper comes the *production phase*, during which simulation averages can be computed. Finally, the simulation is analysed.

2.4.3.1 Assigning Initial Values

The initial configuration of the system can often affect the results of a simulation, and thus it needs to be selected with care. For simulations of systems at equilibrium it is important to choose a starting configuration close to the structure that we are aiming to simulate. The initial coordinates {r(0)} can be obtained from experiments such as X-ray crystallography and NMR measurements, from

computer models generated by a variety of modelling techniques or from a combination of the two. However, some preparation steps are often required before most structures can be used as starting configurations in a MD simulation. For instance, the coordinates for hydrogen atoms are often not available in structures determined by X-ray. The addition of these coordinates to the initial structure is required to initiate the simulation. The available algorithms for assigning protonation states generally lie in two categories. *Geometric methods* place hydrogen atoms based upon optimization of hydrogen bond networks, whereas *electrostatic methods* determine the hydrogen coordinates based upon electrostatic effects considerations and often require the calculation of pKa values. The Protonate3D methodology in Molecular Operating Environment (MOE)[117] adds protons to macromolecules by selecting a protonation state for each chemical group with minimum total free energy of the system, and the results show good overall agreement with high resolution X-ray structures. With the provision of all the coordinates, the initial structure can be further refined by energy minimization in order to relieve local stresses caused by nonbonded overlaps, and to relax bond length and bond angle distortions in the experimental structure. These stresses often result from the empirical nature of the energy function and the average nature of the experimentally obtained structures.

For a molecular dynamics simulation, initial velocities (v_x , v_y , v_z) may be chosen randomly from the Maxwellian velocity distribution at a temperature $\boldsymbol{T}$:

$$\boldsymbol{P}(v)dv = \left(\frac{m}{2\pi k_B \boldsymbol{T}}\right)^{1/2} \exp\left[\frac{-mv^2}{2k_B \boldsymbol{T}}\right] dv \qquad (2\text{-}38)$$

This initial assignment is not at equilibrium; moreover, high velocities may be accidentally assigned to a localized cluster of atoms, leading to unstable

simulation. It is good practice to initiate a simulation by heating up the system step-by-step to avoid this problem. Velocities are first assigned at a low temperature, which is then gradually increased until the simulation reaches the desired temperature. We can use equipartition theorem to define an instantaneous temperature, $\boldsymbol{T}(t)$, at time t in terms of the mean kinetic energy by:

$$\boldsymbol{T}(t) = \frac{1}{k_B N_{dof}} \sum_{i=1}^{N_{dof}} m_i |v_i|^2 \tag{2-39}$$

where N_{dof} is the number of unconstrained degrees of freedom in the system ($N_{dof} = 3N - n$, where N is the number of atoms and n is the number of constraints). We can adjust a mean kinetic energy to match a desired temperature $\boldsymbol{T}_0$ by scaling the velocities with a factor of $[\boldsymbol{T}_0/\boldsymbol{T}(t)]^{½}$.

2.4.3.2 Selecting the Integration Time Step

The size of the time step Δt greatly affects the magnitude of the error in integration algorithms. A small time step means better integration quality; however, it also means that more integration steps are needed to cover the same amount of phase space. There is a trade-off between simulating the correct trajectory and covering the phase space. In general, it is advantageous to use a large time step that will give accurate dynamics. An appropriate time step should be approximately one-tenth the time of the highest frequency motion in the modelled system.

Table 2-1. Suggested time steps for the different types of motion in various systems. Adapted from reference 89.

System	Types of motion	Suggested time step (s)
Atoms	Translation	10^{-14}
Rigid molecules	Translation and rotation	5×10^{-15}
Flexible molecules, rigid bonds	Translation, rotation, torsion	2×10^{-15}
Flexible molecules and bonds	Translation, rotation, torsion, vibration	10^{-15} or 5×10^{-16}

The typical motions together with suggested time steps in various systems are shown in Table 2-1, which can be used to choose an appropriate time step. For flexible molecules, the fastest motions are the stretching vibrations of the bonds between hydrogen atoms and heavy atoms with frequencies about 3000 cm^{-1} (10 fs $= 1 \times 10^{-14}$ s). This extremely small time step limits the time scale of molecular dynamics simulations.

Sophisticated algorithms that allow use of longer time steps to enable longer simulations have been developed. A basic rationale is to remove the fastest motions from the numerical integration and treat them properly in some other way, such that the next fastest molecular motion can be used to decide the time step. The most commonly used algorithm to account for the hydrogen coordinates is SHAKE.[118] In the SHAKE algorithm, the hydrogen coordinates have been adjusted iteratively until the deviation of the current length $d_k(t)$ of the kth bond from its reference bond length d_k^0 satisfied to within a given tolerance value ε:

$$s_k = \left[d_k(t)^2 - d_k^{0\,2}\right] / d_k^{0\,2} < \epsilon \qquad (2\text{-}40)$$

In other words, SHAKE uses an iterative scheme that adjusts the coordinates of hydrogen atoms after each integration step in order to simultaneously satisfy all the constraints. This procedure may be iterated until s_k is smaller than ε for all values of k. It is most useful when we simulate large molecules.

2.4.3.3 Stability of Integration

When a simulation has picked up errors along the way, it becomes unstable. It is therefore important to maintain the accuracy and stability of the simulation for the duration of interest. The integration error will either cause an unstable simulation

to crash, if one is lucky, or continues till its designated termination but give inaccurate trajectories that deviate considerably from true trajectories.

Quantities that are expected to be conserved during the simulation can be used to measure the stability of the simulation. To be considered conserved, an adequate fluctuation ratio for the total energy, ***E*** (the **H**amiltonian) in Equation (2-26) is:

$$\frac{\Delta E}{E} < 10^{-4} \quad \text{or} \quad \log_{10}\left(\frac{\Delta E}{E}\right) < -4 \tag{2-41}$$

The total linear momentum ***P*** and the total angular momentum ***L*** also become conserved after equilibration. However, for simulation schemes with constant adjustments of velocities, momentum conservation can no longer be used for probing the stability of the simulation. Furthermore, Newton's equation is time reversible. This means it should be able to retrace back its original trajectory in phase space (when the sign of the time step Δt is changed to $-\Delta t$). However, numerical errors will cause any two trajectories that are initially very close to diverge from one another although it is still possible to keep a stable integration that exhibits time reversibility for short periods of time.

2.4.3.4 Boundaries

In order to simulate macroscopic properties, it is essential to apply appropriate boundary conditions that mimic the presence of an infinite bulk surrounding the model system and minimise surface effects. This can be achieved by *periodic boundary conditions*. Imagine a cubic box of particles which is replicated infinitely in all directions such that the central box in three dimensions is surrounded by 26 identical cubes, which are again surrounded by 98 boxes, etc. The coordinates of the particles in the periodic image boxes can be obtained simply by adding or subtracting integral multiples of the box sides. If a particle

leaves the central box through the right wall during the simulation, then its image will enter the box through the left wall from the neighbouring box. Thus, the number of particles in the central box remains constant.

Periodic boundary conditions are widely applied in computer simulations, but we should always be aware that they may lead to spurious correlations not present in a truly bulk system. For systems where longer-range electrostatic interactions play important roles, it is important to use a cutoff distance less than half the periodic box length to ensure that a given particle interacts only with the nearest periodic image of any other particles.

Alternatively, the simplest way to incorporate explicitly solvent molecules in simulations is to surround the molecule with a 'skin' of solvent molecules. If the skin is sufficiently deep, it is equivalent to model a solute molecule inside a 'drop' of solvent. The number of solvent molecules required in this approach is usually significantly fewer compared to the analogous periodic boundary simulation. Boundary effects can be transferred from the molecule-vacuum interface to the solvent-vacuum interface such that a realistic treatment of the solute may be achieved.

2.4.3.5 Long-Range Forces

As mentioned in the foregoing section, the charge-charge interaction, which decays as r^{-1}, will extend beyond the boundary of a box and thus is particularly problematic in molecular simulations. It is important to properly model these long-range forces when simulating conformations of biological molecules such as polynucleotide chains. Among a variety of methods developed to handle long-range forces, the Ewald summation[119] has been used successfully to liquid

simulations and is a standard for macromolecular simulations.[113] In this method, a particle interacts with all the other particles in the central simulation cell and with all of their images in an infinite array of periodic cubes.

The idea of the Ewald summation is to split the interaction, which converges extremely slowly, into two series, each of which converges much more rapidly. The first contribution is obtained by surrounding each point charge with a neutralising charge distribution of equal magnitude but of opposite sign centred at the position of the point charge. The conversion of the sum over point charges to a new sum of the interactions between the charges plus the neutralising distributions leads to zero net charge. The interaction between the screened point charges is short-ranged and converges very rapidly. A smooth charge distribution is now added to the system which exactly offset the first neutralising distribution. The second summation can be evaluated in reciprocal space, and also converges much more rapidly than the original point-charge sum. Nevertheless, one needs to bear in mind that the Ewald method is not without problems. It tends to reinforce artefacts resulting from imposing periodic boundary conditions and is computationally rather expensive to implement.

2.4.3.6 Simulation Protocol

Every molecular dynamics simulation comprises several steps. A simulation protocol refers to these steps all together. The main steps common to most MD simulation protocols are: (i) preparation of the initial configurations such as addition of hydrogen atoms, energy minimization, and specification of initial velocities, (ii) gradual increase on the velocities to the desired temperature, (iii) a relatively long dynamic simulation that may take from tens of picoseconds to

several hundred picoseconds to bring the system to equilibrium from the starting configuration and to ensure that the simulation is stable, (iv) the dynamic simulation after the equilibration phase called production, where the trajectory generated can be stored for further analysis, and (v) analysis of the resulting trajectories.

2.4.4 Molecular Dynamics in Various Ensembles

The original MD technique was performed in the constant NVE ensemble. It is important to perform simulations in other ensembles relevant to experimental conditions, such as the canonical (constant-NVT) ensemble associated with a Helmholtz free energy or the isothermal-isobaric (constant-NPT) ensemble associated with a Gibbs free energy. Herein, I outline the most widely used methods of MD simulations under conditions of constant temperature or constant pressure.

2.4.4.1 Constant Temperature MD

An obvious approach to maintain the temperature of a system during an MD simulation is to rescale the velocities at each time step by a factor of $[\boldsymbol{T}_0/\boldsymbol{T}(t)]^{½}$, where $\boldsymbol{T}(t)$ is the temperature at time t defined in Equation (2-39) and $\boldsymbol{T}_0$ is the desired temperature. A further refinement of this approach is to couple the system to a heat bath at a temperature $\boldsymbol{T}_0$ using velocity scaling.[120] The bath can supply or remove heat from the system like a source of thermal energy, such that the velocities can be rescaled at each step and the rate of change of temperature is proportional to the difference in temperature between the bath and the system:

$$\frac{d\boldsymbol{T}(t)}{dt} = \frac{1}{\tau}(\boldsymbol{T}_0 - \boldsymbol{T}(t)) \qquad (2\text{-}42)$$

where τ is a coupling parameter, which determines the strength of the coupling to the heat bath. The change in temperature between successive time steps can be expressed as:

$$\Delta \boldsymbol{T} = \frac{\delta t}{\tau}(\boldsymbol{T}_0 - \boldsymbol{T}(t)) \tag{2-43}$$

The velocities can thus be scaled by a factor λ:

$$\lambda = \left[1 + \frac{\delta t}{\tau}\left(\frac{\boldsymbol{T}_0}{\boldsymbol{T}(t)} - 1\right)\right]^{1/2} \tag{2-44}$$

If τ is large, the coupling is weak. With the decreasing τ, the coupling will be stronger. When the coupling parameter equals the time step ($\tau = \delta t$) the algorithm is equivalent to the simple velocity scaling method. For the time step of 1 fs, the coupling parameter of 0.4 ps ($\delta t/\tau \approx 0.0025$) is suggested as an appropriate value. However, these two methods do not generate rigorous canonical averages and do not correspond to exact canonical ensembles.

In the stochastic collisions method, the molecules of the system interact with the particles of a heat bath at a specified temperature.[121] A particle is randomly selected at intervals and its velocity is obtained from the Maxwell-Boltzmann distribution in Equation (2-38) corresponding to the desired temperature $\boldsymbol{T}_0$. The mean frequency with which each particle experiences a stochastic collision, ν, chosen from a Poisson distribution:

$$\boldsymbol{P}(\nu, t) = \nu \exp(-\nu t) \tag{2-45}$$

where $\boldsymbol{P}(\nu, t)$ is the probability that the next collision will take place in the time interval δt.

Another approach to isothermal MD simulation is to treat the dynamics of a system in contact with a thermal reservoir by considering the thermal reservoir as

an integral part of the system and introducing an additional degree of freedom, denoted *s*, that represents that reservoir.[122,123] In a simulation of this extended system, energy is allowed to flow back and forth from the reservoir to the system, and each state of the extended system corresponds to a unique state of the real system. The velocities of the atoms in the real system are related to the time-derivatives of position by:

$$\mathbf{v}_i = s\frac{dr_i}{dt} \tag{2-46}$$

where r_i is the position of particle *i* in the simulation and $\mathbf{v}_i$ is considered to be the velocity of the particles in the real system. The time step $\delta t'$ is associated with the time step in real system δt by:

$$\delta t = s\delta t' \tag{2-47}$$

The reservoir has potential energy $(f + 1)k_BT \ln s$, where *f* is the number of degrees of freedom (3N − 3 if the total momentum is fixed) and *T* is the desired temperature. The kinetics of the reservoir can be expressed as $(Q/2)(ds/dt)^2$, where *Q* is the thermal inertia parameter with dimensions of energy × (time)^2 which determines the coupling between the reservoir and the real system and thus controls the temperature fluctuations. The time-scale variable *s*, its conjugate momentum p_s and the parameter *Q* are introduced to express the Hamiltonian of the extended system of the N particles plus additional degrees of freedom:

$$\boldsymbol{H}_{\text{Nosé}} = \sum_{i=1}^{N}\frac{P_i^2}{2m_i s^2} + U(q) + \frac{P_s^2}{2Q} + (f+1)k_BT\ln s \tag{2-48}$$

The time step in real time can fluctuate as the value of the additional degree of freedom *s* changes. To avoid fluctuating time intervals in the simulation, a slightly

different set of equations have been derived with no involvement of the time-scaling parameter *s*. During the simulation, the following quantity is conserved:

$$\bar{\boldsymbol{H}}_{\text{Nosé}} = \sum_{i=1}^{N} \frac{P_i^2}{2m_i} + U(q) + \frac{\xi^2 Q}{2} + f k_B T \ln s \qquad (2\text{-}49)$$

where ξ (= p_s/Q.) is the thermodynamic friction coefficient. It should be stressed, however, that this $\bar{\boldsymbol{H}}_{\text{Nosé}}$ is no longer a Hamiltonian. The adjustable parameter *Q* regulates the energy flow between the system and the reservoir. A small value of *Q* corresponds to a low thermal inertia of the bath and results in poor equilibration and rapid temperature fluctuations. However, a large value of *Q* results in slow energy exchange, and in the limit of infinite *Q*, conventional MD is regained. It may be necessary to choose *Q* by performing a series of trial simulations for a test system in order to observe how well the desired temperature is maintained.

2.4.4.2 Constant Pressure MD

Many experiments are performed under conditions of constant temperature and pressure, and so simulations in the isothermal-isobaric (NPT) ensemble are most directly relevant to experimental data. For MD simulations at constant pressure the volume is considered as a dynamical variable that changes during the simulation. Many schemes used for pressure control are analogous to those used for temperature control. The pressure can be kept constant by simply scaling the volume. The rate of change of pressure can be expressed as:

$$\frac{d\boldsymbol{P}(t)}{dt} = \frac{1}{\tau_p}(\boldsymbol{P}_0 - \boldsymbol{P}(t)) \qquad (2\text{-}50)$$

where τ_p is the coupling constant, $\boldsymbol{P}_0$ is the pressure of the bath, and $\boldsymbol{P}(t)$ is the pressure at time t. The volume of the simulation box is scaled by a factor λ:

$$\lambda = \left[1 - \frac{\delta t}{\tau_p}(\boldsymbol{P}_0 - \boldsymbol{P}(t))\right] \quad (2\text{-}51)$$

This is equivalent to scaling the atomic coordinates by a factor $\lambda^{1/3}$, yielding the new positions:

$$r_i' = \lambda^{1/3} r_i \quad (2\text{-}52)$$

This method does not correspond to any ensembles, and therefore it is not possible to determine the meaning of fluctuations in any observed quantity.

Alternatively, the system can be coupled to a 'pressure bath', analogous to a temperature bath.[121] An extra degree of freedom *V*, corresponding to the volume of the box, is added to the system. This coupling can be considered as the action of a piston on a real system. The piston is associated with a kinetic energy, $\frac{1}{2}Q(dV/dt)^2$, where Q is the 'mass' of the piston. The potential energy of the piston is *PV*, where *P* is the desired pressure and *V* is the volume of the system. A piston of small mass leads to rapid oscillations in the box size. A large mass results in slow exploration of volume-space, and an infinite mass returns normal MD. The extended-system temperature-scaling method employs a scaled time, and likewise, in the extended pressure method the coordinates of the extended system, r_i', are related to the real coordinates r_i by:

$$r_i' = V^{-1/3} r_i \quad (2\text{-}53)$$

2.5 Quantum Mechanical Calculations

2.5.1 Computational Details for the Study on Photochemistry of Pt(IV) Anticancer Agents

The geometries of molecules were optimized with the Amsterdam Density Functional 2007 program (ADF)[124] at the gradient-corrected DFT level using the BP86 functional in combination with the uncontracted triple-ζ+polarisation (TZP) STO basis sets. A small frozen core was used for efficient treatment of the inner atomic shells. Relativistic effects were considered by the zeroth-order regular approximation (ZORA)[58,59]. Vibrational frequency calculations confirmed that the structures reported are minima on the potential energy surfaces. Ziegler and Rauk's energy decomposition scheme has been used to study the metal-ligand bond for selected structures.[125-128] The lowest fifty singlet electronic excitation energies and their corresponding oscillator strengths were calculated by the statistical average of orbital potentials (SAOP)[60] in combination with a TZ2P basis sets (i.e. TZP with an additional polarization function) using scalar relativistic Time-Dependent Density Functional Theory (TDDFT) as implemented in the ADF program. Spin-orbit coupling was evaluated in the Pt(IV) complexes to enhance the comparison with available experimental data as previous theoretical study has shown that relativistic effects can have a significant effect on the simulated electronic spectra of heavy metal complexes.[129] The results show that the main feature of the spectra is similar with or without spin-orbit coupling. Electronic transitions were also computed using TDDFT with the Conductor like Screening Model (COSMO) for *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$ at the previously optimized structure in gas phase. COSMO is an approximate approach to include solvent effects for reactions in solution. In COSMO the solvent is treated as a dielectric continuum, surrounding the molecular cavity of the solute molecules. A rigid sphere of the solvent molecules is used to determine how far the solvent can penetrate into the region occupied by the solute molecules, the

solvent accessible surface of the cavity. The overall surface is partitioned into small surface triangles and each triangle is assigned an adjustable point charge, which are determined from the charge density and corresponding potential of the molecule itself in the SCF procedure and by assuming an ideal conductor where the electric potential on the surface disappears. Thus, along with the computed charge distribution of the molecule, the energy of the interaction between the solvent and the solute molecule can be calculated. The addition of solvent effects does not show improvement in terms of the absorption maxima compared to the experimental data (Table A1-2, Appendix 1). Thus, all subsequent calculations were carried out using scalar relativistic TDDFT at a relatively low computational cost and no solvent effects were accounted. The geometry optimization of the lowest-lying triplet state of complex **11** was performed using a spin-unrestricted BP86/TZP method with a net total spin polarization chosen to be 2.

The simulated UV spectra were produced by the summation of Gaussian peaks, generated for each of the calculated absorption maxima according to:

$$I = Ae^{-(x-b)^2/(2c^2)} \quad (2\text{-}54)$$

where I is the peak intensity as a function of wavelength (x), A is the calculated absorption maximum, b is the wavelength at which the calculated maximum occurs and c is the peak width parameter related to the full width at half maximum (FWHM) resolution by:

$$FWHM = 2c\left(\sqrt{2\ln 2}\right) \quad (2\text{-}55)$$

The model was devised to overcome issues inherent in the comparison of computed "point spectra" with those generated by experiment. It accounts for coalescence of neighbouring peaks, allowing this to be visualised as a function of

peak width, but takes no account of dynamic peak width variations arising from the different types of electronic states involved in each transition.

2.5.2 Quantum Mechanical Calculations for New LFMM/MMFF94 and LFMM/AMBER94 Charge Schemes

To estimate partial charges of Pt-guanine complexes, the geometry of a model Pt-guanine complex was fully optimized using the ADF 2007 program[57] with the BP86 functional[130,131] in conjunction with the a triple-ζ plus double polarization basis set (TZ2P) on all atoms and a scalar ZORA[59] relativistic and dispersion[132] correction. To maintain compatibility with our previous parameterization, CHelpG[133] charges were computed using the previously optimized structure at the HF/6-31G(d) level using the Gaussian 03 program[61] with the LanL2DZ[134] basis set with an extra set of f functions using the exponents determined by Frenking et al. for the metal[135] with Pt given a van der Waals radius of 1.8 Å .

2.6 Analysis Techniques

The value of any computer simulation lies in the quality of the information extracted from it. The analysis procedure needs to be contemplated before starting the simulation, as the goals of the analysis will often determine the general character of the simulation. Herein, I give an overview of the techniques that have been used to analyse the MD trajectories presented in this thesis.

2.6.1 Average Structure

To compare with experimental and/or different MD structures, average MD structures of DNA were calculated as below. All *x, y,* and *z* coordinates of heavy (non-hydrogen) atoms were read in from the specified section of the trajectory.

Translation and rotation motions were removed by least-squares fitting to a reference structure (usually the initial structure at t = 0). The average x, y, and z coordinates were then calculated using Equation (2-56).

$$average(x, y, z) = \frac{\sum_{i=1}^{N_S}((x_i, y_i, z_i)}{N_S} \quad (2\text{-}56)$$

where N_s is the number of snapshots in the trajectory.

2.6.2 Root-Mean-Square Deviation

In order to compare different simulations to one another, or to compare the simulation to experimental results, we need to compute average values and fluctuations. The most common mathematical expression of a molecule's motion within a trajectory is the root-mean-square deviation (RMSD) of the atoms in the system with respect to a reference structure. For a stable system, the RMSD is expected to fluctuate about a constant value. The translational and rotational motions are first removed by least-squares fitting to a reference structure. The RMSD is then calculated by Equation (2-57).

$$\mathrm{RMSD}_{(\mathrm{i,j})} = \left[\frac{1}{N_s}\sum_{k=1}^{N_s}(x_{ik} - x_{jk})^2 + (y_{ik} - y_{jk})^2 + (z_{ik} - z_{jk})^2\right]^{1/2} \quad (2\text{-}57)$$

2.6.3 DNA Backbone Conformation Parameters

The IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) has made recommendations for abbreviations and symbols to describe conformations of polynucleotide chains. The backbone of a polynucleotide chain comprises a repeating unit of six single bonds, P–O5', O5'–C5', C5'–C4', C4'–C3', C3'–O3' and O3'–P. The torsion angles related to these bonds are indicated by the symbols α, β,

γ, δ, ε, ζ, respectively. The symbols α(*i*)–ζ (*i*) are referred to torsion angles of bonds within the i^{th} nucleotide unit as illutrated in Figure 2-6 and 2-7.[136]

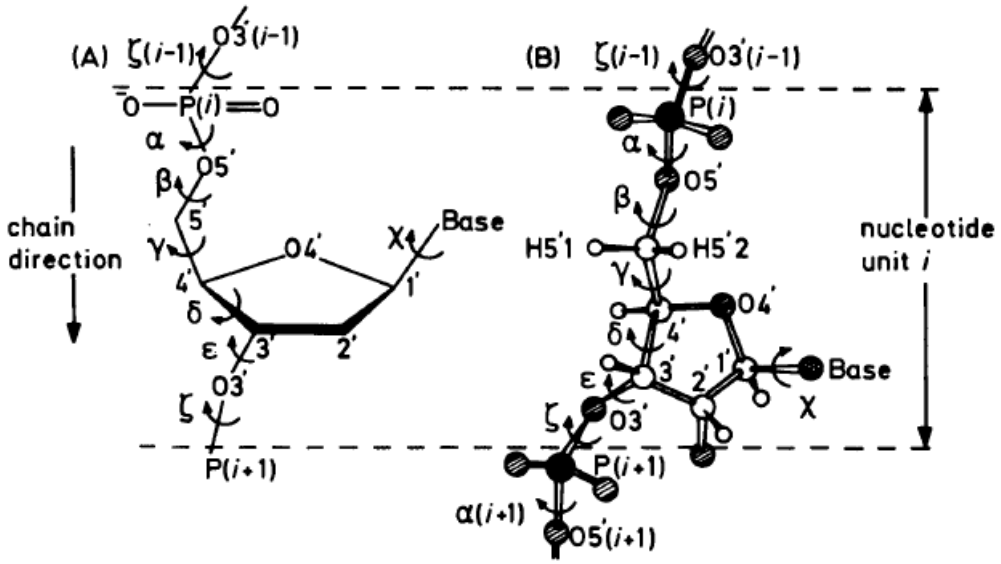


Figure 2-6. Section of a polynucleotide backbone showing the atom numbering and the notation for torsion angles. Reproduced from reference 137.

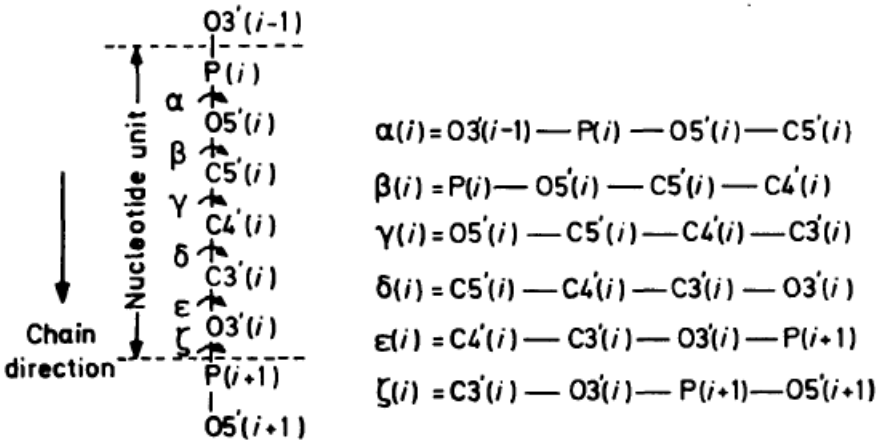


Figure 2-7. Torsion angles for backbone conformations of the *i*th nucleotide in polynucleotide chains. Reproduced from reference 137.

2.6.4 DNA Helicoidal Parameters

Curves+[138] is a conformational analysis program for nucleic acids with up to four strands and with either standard or chemically modified bases and backbones.

Compared to the original Curves algorithm widely used for nucleic acids structural studies, this program is algorithmically simpler and computationally much faster. In addition, it provides helical and backbone parameters and a full analysis of groove widths and depths. Diagrammatic definitions of parameters used to describe the geometry of complementary base pairs and sequential base pair steps are presented in Figure 2-8 and 2-9. The intra-base pair parameters include three translations, shear, stretch and stagger, and three rotations, buckle, propeller and opening. Likewise, the inter base pair parameters contains three translations, shift, slide and rise, and three rotations, tilt, roll and twist. Therefore, Curves+ is used to analyse the local and global structural parameters for nucleic acids in Chapter 5.

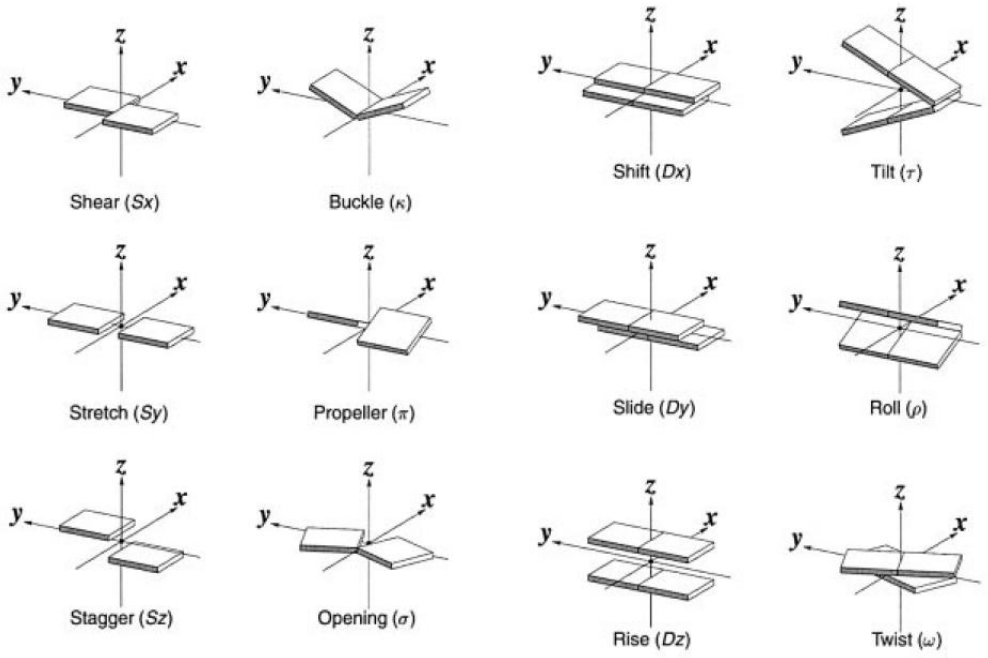


Figure 2-8. Pictorial definitions of the intra-base pair (left) and inter-base pair (right) parameters in Curves+. Images illustrate positive values of the designated parameters. Diagram adapted from reference 139.

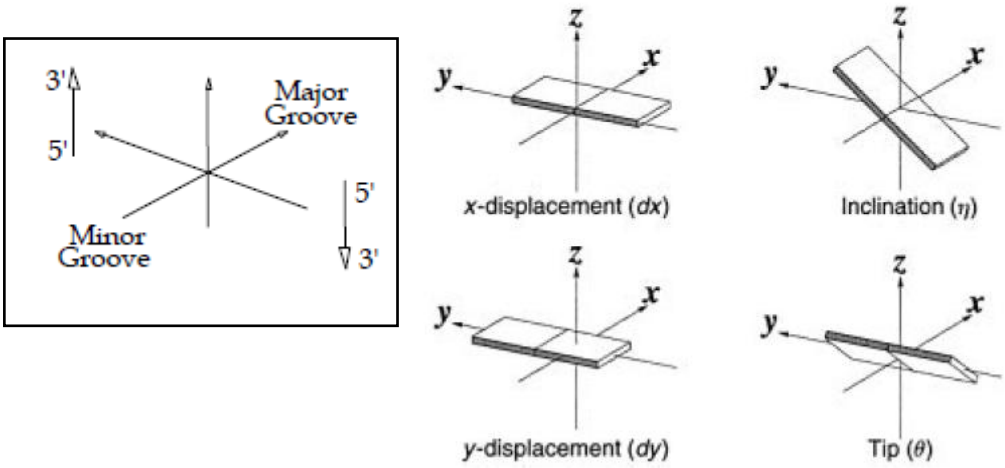


Figure 2-9. Illustration of the base pair reference frame (left) and of the local helical frame (right). Images demonstrate positive values of the designated parameters. Diagram adapted from reference 139.

2.6.5 Structural Forms of DNA

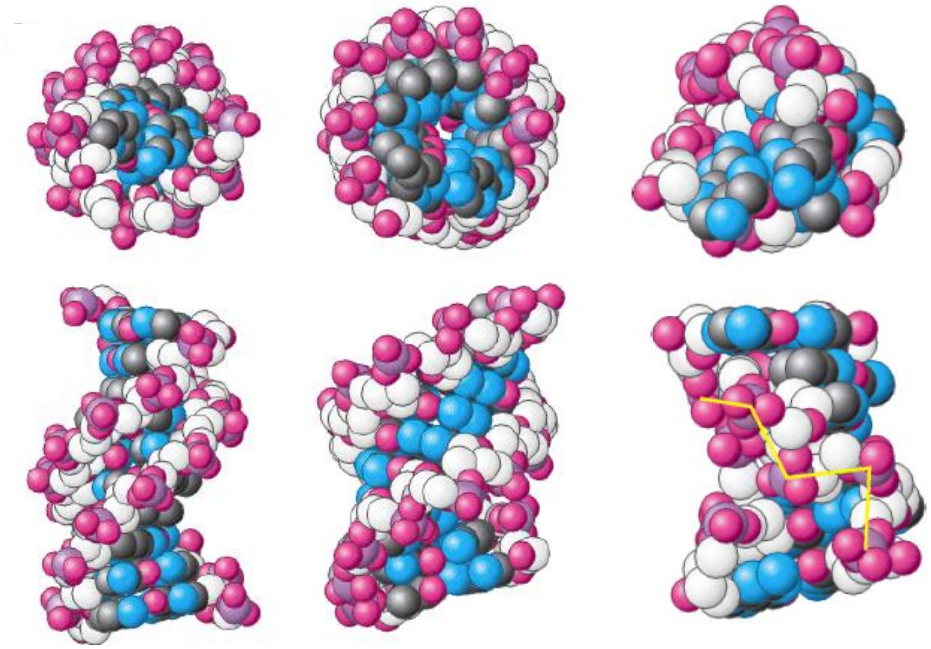


Figure 2-10. Structures of A-, B-, and Z-forms of DNA. Adapted from reference 140.

The classical model of double stranded DNA deduced by Watson and Crick is known as B-DNA helix. Due to the flexibility of its backbone, DNA can assume a

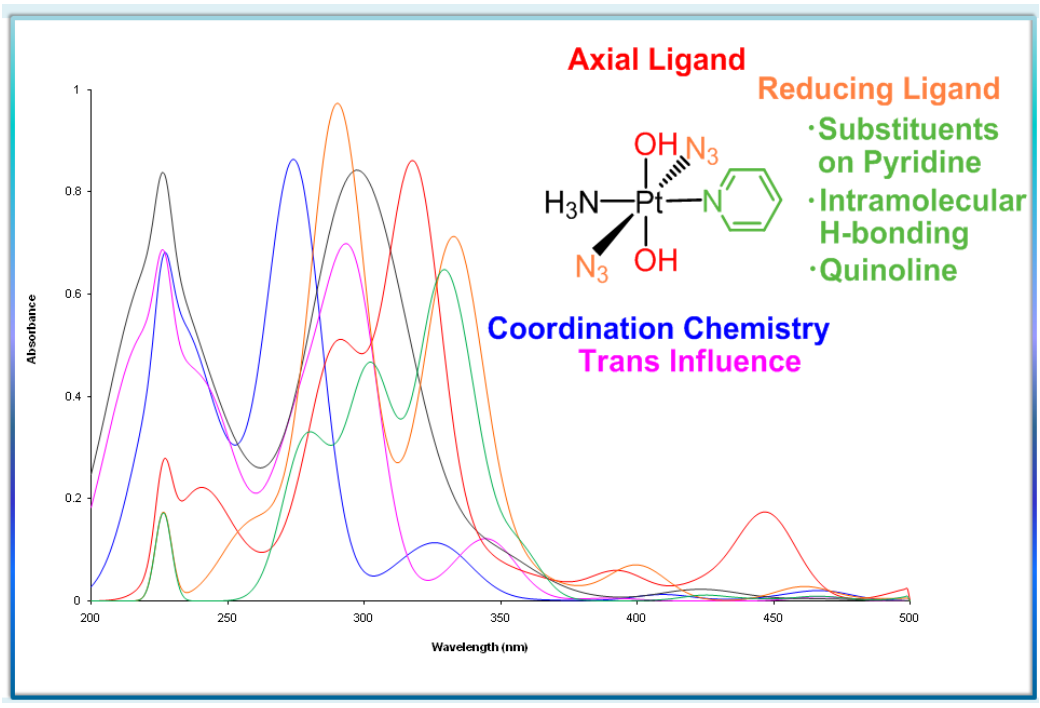
variety of structural forms. Along with B-DNA two other forms, A-DNA and Z-DNA, are biologically significant (Figure 2-10). The main structural parameters for standard A-, B-, and Z-forms of DNA are provided in Table 2-2.

Table 2-2. Summary of the main structural parameters for standard A-, B-, and Z-forms of DNA.[141]

Parameter	**A-DNA**	**B-DNA**	**Z-DNA**
Shape	Broadest	Intermediate	Narrowest
Helix sense	Right	Right	Left
α	−85°	−47°	60°/160°
β	−152°	−146°	−175°/−135°
γ	46°	36°	178°/57°
δ	83°	156°	140°/95°
ε	178°	155°	−95°/−110°
ζ	−46°	−95°	−35°/−85°
Sugar pucker	C3'-endo	C2'-endo	C3'-endo/ C2'-endo
Glycosidic bond	*anti*	*anti*	Alternating *anti* and *syn*
Base roll	12°	0°	1°
Base tilt	20°	5°	9°
Base twist	32°	36°	11°/50°
Helical diameter	25.5 Å	23.7 Å	18.4 Å
Bases/turn of helix	11	10.4	12
Rise/base pair	2.3 Å	3.4 Å	3.8 Å
Major groove	Narrow, deep	Wide, deep	Flat
Minor groove	Broad, shallow	Narrow, deep	Narrow, deep

Chapter 3

A Computational Approach to Tuning the Photochemistry of Pt(IV) Anticancer Agents



3.1 Introduction

Since the serendipitous discovery of the inhibition of cell division by cisplatin in the 1960s, platinum-based drugs have been widely used against various tumours.[1,5] Much effort has been devoted to the design of novel Pt anticancer drugs to make chemotherapy safer for patients, in particular, lessening severe side effects,[142,143] increasing oral bioavailability,[144,145] and overcoming drug resistance.[146-151] Platinum-based drugs disrupt DNA synthesis by mechanisms common to all actively dividing cells, regardless of whether they are cancerous or not, which generates severe side effects. In addition, the large volume of damaged cells further cause altered cellular transport, enhanced repair of distorted DNA, overexpression of DNA-damage recognition proteins and decreased apoptosis contributing to tumour resistance to Pt drugs.[6,19] Thus, the design of drugs with increased selectivity is of great importance for significant clinical advantages over the current drugs.

Prodrugs to improve therapeutic targeting and controlled release at the tumour site show promising possibilities for increased selectivity.[152] Improved delivery of platinum drugs with liposomal-based or polymer-based vehicles is an attractive strategy and has been in clinical trials.[22-24] However, creating delivery systems that target specific sites to achieve good release and activation is challenging. On the other hand, a photosensitizing agent in combination with light provides a prominent avenue to achieve accurate targeting.[25,26]

An approach to platinum based photochemotherapy of cancer is to design photoactive Pt complexes which have a strong absorbance at longer wavelengths, increased selectivity, and facile metabolism[25,153]. Platinum(IV) complexes are

much more inert to reaction *in vivo* but readily undergo photoinduced reduction and ligand substitution. Moreover, photoreduction reactions from excited states are more reactive with macromolecules and, therefore, can be used effectively for triggering specific interactions between the metal complex and target biomolecules (such as DNA, RNA or proteins).[33-35,154-159] Altering the electronic configuration by irradiation can lead to remarkable differences in chemistry, including the effects of changes in geometry, dipole moment, redox characteristics and acid-base properties, and can bring about reactions that normally do not occur under thermal conditions.[27] Exploration of the factors that control the photochemistry of Pt(IV) complexes may thus allow the design of effective anticancer agents.

As irradiation with longer wavelength helps reduce cell damage and deeper penetration can be achieved,[160] the aim of this work is to achieve photoactivation of Pt(IV) complexes at longer wavelengths. *Trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(py)]$ (**1**) is highly phototoxic and, in particular, pyridine is suggested to contribute to the potency via a different mechanism of action compared to cisplatin.[35] Therefore, complex **1** has been used as the starting point and the template for our theoretical modelling. By tuning *cis*/*trans* geometries, *trans* influence, the OR^- (R=H, Ac) ligands, (pseudo)halogen (N_3^-, I^-) ligands, electron-withdrawing/donating/delocalizing substituents on pyridine ligands, and intramolecular H-bond, the effects of the coordination environment on the electronic excitations of Pt(IV) complexes have been explored (Figure 3-1). The present study has also evaluated the influence of various chemical modifications on (i) the electronic excitations of the intense absorption band (normally at shorter wavelength) and of low-energy absorption band (the "tail" normally at longer

wavelength), (ii) the molecular orbitals involved in the electronic transitions in the UV-Vis bands, and (iii) the nature of bonding. In addition, *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] has been synthesised and characterised by Yao Zhao in order to compare its absorption properties with the calculations. A theoretical study of ground and excited state properties of photoactive Pt(IV) diazido was reported while this thesis was in preparation.[161] Our results on the replacement of OH^- ligand by aromatic OR^- and the formation of azidyl radicals N_3• upon photoreduction largely agree with those reported in this work, but our present study focuses on a systematic survey of various chemical influences on the photochemistry of Pt(IV) complexes in intense band and low-energy absorption in order to provide guidelines for future synthesis of this class of compounds. Relevant computational details were given in Chapter 2.

3.2 Influence of Various Chemical Modifications

The electronic excitations of Pt(IV) complexes are evaluated by the influence of various chemical modifications with reference to complex **1** (see Figure 3-1). Calculated excitation energies for the intense band are in good agreement with experimental measurements with errors less than 0.3 eV. Electronic *trans* influences are well-understood ground-state phenomena,[103] but relatively little is known about their effects on photophysical properties. OR^- and halogen ligands have a large effect on the rate of reduction,[162] and thus it is interesting to study their influence on electronic transitions in Pt(IV) complexes. As earlier reports[25,162,163] showed that Pt(IV) complexes with OAc^- or halogen ligands are often easily reduced in cells and are less likely to have a noticeable difference between light and dark toxicity, in this study the design with *trans* OH^- and *trans*

N_3^- groups is preferred. Pyridine with its range of possible substituents is a useful ligand for structure-property studies, including the photochemistry of Pt(IV) compounds. All these unique chemical properties lead us to the current construction of the theoretical models in this work. Figure 3-3 summarises the key features of the computed intense and lowest energy absorptions under the influence of various chemical modifications given above. Figure 3-4 displays simulated UV spectra of selected Pt(IV) complexes.

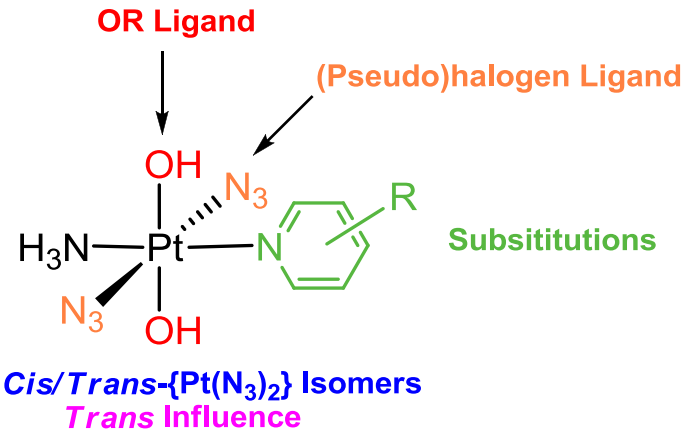


Figure 3-1. Various chemical modifications to explore the photochemistry of a potent photocytotoxic platinum complex, *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(pyridine)]$ (**1**). Changes in the coordination geometry and *trans* influence were considered. Substitutions (R) on the imine ligands, OR and (pseudo)halogen ligands were also included in the calculation. The substitutions on the imine ligands vary from withdrawing/donating groups/quinoline to intramolecular H-bonding involving OH ligand

3.2.1 Comparison between Experimental and Calculated Structures and Absorption.

The fully optimized structures of *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$, and its all *trans* isomer show that bond lengths and angles are in good agreement with the respective experimental values, with a root-mean-square deviation (RMSD) of

0.046 Å for Pt−L bonds and 0.9° for L−Pt−L angles (Table A1-1, Appendix 1).[34,164] The computed absorption maximum for the complex *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)(py)]$ (**2**) (py = pyridine) (273 nm, Table 3-2) is 20 nm shorter than that of its all *trans* isomer complex **1** (293 nm, Table 3-1). The same trend was observed for the *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$ (271 nm, Table 3-1) and *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$ (295 nm, Table 3-1), which is consistent with the previously reported experimental data, 256 nm and 285 nm respectively (Table 3-1).[34] The comparison of the simulated and experimental UV-Vis spectra reveals smaller discrepancy in the intense band (λ_{max}) for complex **1** and *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(quin)]$ (**13**) (quin = quinoline) (Table 3-1 and Figure 3-2) compared to *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$ and *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$. The high electron density of pyridine and quinoline may increase the accuracy of the modelled frontier orbitals.[165] In addition to the electronic excitation energies, the computed intensity of absorption is consistent with the UV-Vis spectrum.[34,164]

Table 3-1. Experimental and calculated absorption properties (intense band) of *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$, *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$, *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(py)]$ (**1**), and *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(quin)]$ (**13**)

λ_{max}, nm ε (M^{-1} cm^{-1})	Tr[a]	Composition	Energy/eV (nm)	Oscillator Strength	Assignment
trans, trans, trans-$[Pt(N_3)_2(OH)_2(NH_3)_2]$					
285 (18900)	11	HOMO−3→LUMO (46%) HOMO−6→LUMO (24%) HOMO−3→LUMO+1 (19%)	4.21 (295)	0.341	LMCT/ LLCT
	19	HOMO→LUMO+3 (67%)	5.24 (237)	0.192	LLCT

cis, *trans*, *cis*-[$Pt(N_3)_2(OH)_2(NH_3)_2$]					
256 (13700)	13	HOMO−5→LUMO+1 (64%) HOMO−4→LUMO (20%)	4.47 (277)	0.134	LMCT/ LLCT
	14	HOMO−4→LUMO+1 (39%) HOMO−5→LUMO (33%)	4.58 (271)	0.156	LMCT/LLCT
trans, *trans*, *trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(py)] **(1)**					
289 (18800)	17	HOMO−6→LUMO (48%) HOMO−3→LUMO (29%) HOMO−3→LUMO+2 (14%)	3.97 (312)	0.119	LMCT/LLCT
	21	HOMO−6→LUMO (46%) HOMO−3→LUMO (37%)	4.23 (293)	0.261	LMCT/LLCT
trans, *trans*, *trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] **(11)**					
287 (18200)	29	HOMO−10→LUMO (68%)	3.99 (311)	0.147	LLCT/ MLCT
	30	HOMO−3→LUMO+1 (59%) HOMO−5→LUMO+1 (13%)	4.08 (304)	0.311	LLCT/LMCT
trans, *trans*, *trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(quin)] **(13)**					
316 (17800) and 294 (28300)	20	HOMO−5→LUMO+2 (31%) HOMO−4→LUMO+2 (17%) HOMO−4→LUMO (15%)	3.98 (311)	0.120	LMCT/LLCT
233 (42600)	45	HOMO−8→LUMO+3 (17%) HOMO−11→LUMO (14%)	5.27 (235)	0.106	LMCT/LLCT/ MLCT/d-d

[a] Tr = transition number as obtained in the TD-DFT calculation.

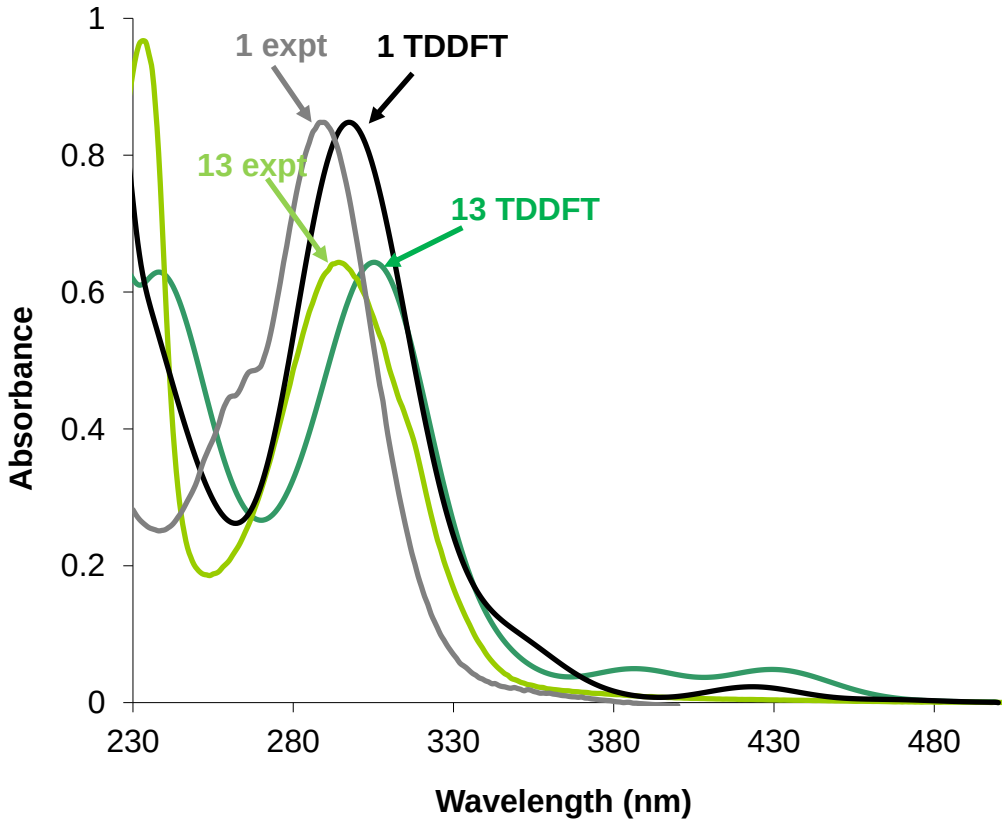


Figure 3-2. Comparison between calculated and experimental UV-Vis spectra for *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(pyridine)]$ (**1**) and *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(quinoline)]$ (**13**). The oscillator strength of complexes **1** and **13** has been scaled by 2.5 and 1.8 times, respectively, for the convenience of comparison.

3.2.2 *Trans* Influence

To investigate the *trans* influence on photophysical properties in Pt(IV) complexes, ligands with strong a *trans* influence were placed *trans* to the N_3^- ligand, the major reducing/leaving ligand. TDDFT predicts intense bands at 295 nm and 306 nm for Me_2S and H^- ligands *trans* to azide in *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)S(CH_3)_2]$ (**3**) and *cis, trans, cis*-$[PtH(N_3)_2(OAc)_2(NH_3)]^-$ (**4**), respectively, significantly longer wavelengths than in complex **2** (273 nm). Comparison of the main absorbance (λ_{max}) for complexes **2, 3** and **4** shows that the

trans influence has a considerable effect on the electronic excitation energies in Pt(IV) compounds (see Figure 3-3 and 3-4 for the comparison with other Pt(IV) complexes).

3.2.3 OR^- (R=H, Ac) and Halogen Ligands

To explore the influence of OR^- and halogen ligands, the electronic transitions in complexes with OH^-**,** OAc^-**,** and I^- ligands were studied. TDDFT assigns an absorption maximum (λ_{max}) at longer wavelength, 318 nm for *trans, trans, trans*-$[Pt(N_3)_2(OAc)_2(NH_3)(py)]$ (**5**) and 320 nm for *trans, trans, trans*-$[Pt(N_3)_2I_2(NH_3)(py)]$ (**6**), compared to complex **1** (293 nm). The excitation energies decrease in the order OH^- > OAc^- ~ I^-. This result indicates that the energy required for a Pt(IV) complex with OH^- to populate an electronic state to the potential energy surface at an energy above the transition states of photoreduction to Pt(II) complexes is higher than the complex with OAc^- or I^-. On the other hand, it was reported that the Standard Hydrogen Electrode adjusted reduction potentials for *cis, trans, cis*-$[Pt(NH_3)_2(X)_2(Cl)_2]$ are in the order of: OH^- > OAc^- > Cl^-.[162,163] It is interesting to note that the excitation energies follow the same trend with the energy difference between the products and reactants in reduction for Pt(IV) complexes with OR^- and halogen ligands. The replacement of the N_3^- ligands in complex **5** by I^- ligands (complex **7** in Figure 3-3) result in the red shift of the intense band at 333 nm. Meanwhile, the "tail" of the absorbance extends from 506 nm in complex **1** to 744 nm in complex **6** and to 732 nm in complex **7**. The trend of the influence on the simulated spectra of the studied complexes is displayed in Figure 3-4.

3.2.4 Substituents on Pyridine, Quinoline, and Intramolecular Hydrogen Bonding

To elucidate the relationship between electronic effects on the pyridine ring and photophysical properties of Pt(IV) complexes, electron-withdrawing (via inductive or resonance effect), electron-donating and electron-delocalizing substituents on pyridine have been considered. OR^- and halogen ligands show a critical influence on electronic transitions. However, the relatively higher energies of the unoccupied frontier orbitals in Pt(IV) complexes with OAc^- or halogen ligands often lead to rapid reduction in cells. The lower excitation energies can be achieved by either increasing the energies of the occupied frontier orbitals or decreasing the energies of the unoccupied frontier orbitals, and thus, the design of photoactivated Pt drugs with greater absorption at longer wavelengths can be better achieved by the latter approach. This concept leads to the exploration of Pt(IV) complexes with an intramolecular hydrogen bond between 2-hydroxypyridine/2-hydroxyquinoline and the hydroxyl group. In general, electron-withdrawing functional groups on pyridine red shift not only the absorption maxima by 13 nm but also the lowest energy absorption by 38–332 nm, while intramolecular hydrogen bonding involving the OH ligands and 2-hydroxyquinoline red shifts the main absorption by 24–32 nm and extends the lowest energy absorption toward the red region by 40–100 nm.

The calculated wavelengths of intense bands for complexes *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)$(2-cyanopyridine)] (**8**), *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)$(2,6-difluoropyridine)] (**9**), and *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)$(2-nitropyridine)] (**10**) are 306 nm, 305 nm and 307 nm, respectively, which reveal that electron-withdrawing by induction and resonance

have a subtle influence on the electronic excitation energies of Pt(IV) compounds. Surprisingly, *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(4-methoxypyridine)] (**12**) shows a maximum at 308 nm, similar to complexes with electron-withdrawing groups on pyridine. Comparison of the predicted bands of complexes **8**, **9**, **10**, and **12** indicates that electron-withdrawing/donating substituents on pyridine slightly extend the main absorption to longer wavelength. In complex **13**, the band centred at 311 nm suggests that stabilization of the unoccupied frontier orbitals by electron delocalization decreases the energies of electronic transitions to a similar extent as electron-withdrawing/donating groups. Moreover, the calculated intense bands of *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(2-hydroxyquinoline)] (**15**) and of *trans, trans, trans*-[$Pt(N_3)_2(OH)_2$(2-hydroxyquinoline)$_2$] (**16**), 317 nm and 325 nm respectively, are red shifted compared to complexes **8**, **9** , **10**, **12** and **13** which shows that intramolecular hydrogen bonding involving OH ligands can have a marked effect. On the other hand, it is interesting to note that the low-energy low-intensity spectroscopic absorption extends from the green region of the spectrum for complex **1** (506 nm) to the red region for complexes **16** (606 nm), **8** (625 nm), and **10** (858 nm), yet is blue shifted by 30 nm for complex **12**. The trend of electronic transitions for the Pt(IV) complexes is showed in Figure 3-3 and 3-4. The results for *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] (**11**) will be analyzed and discussed later in Section 3.5.

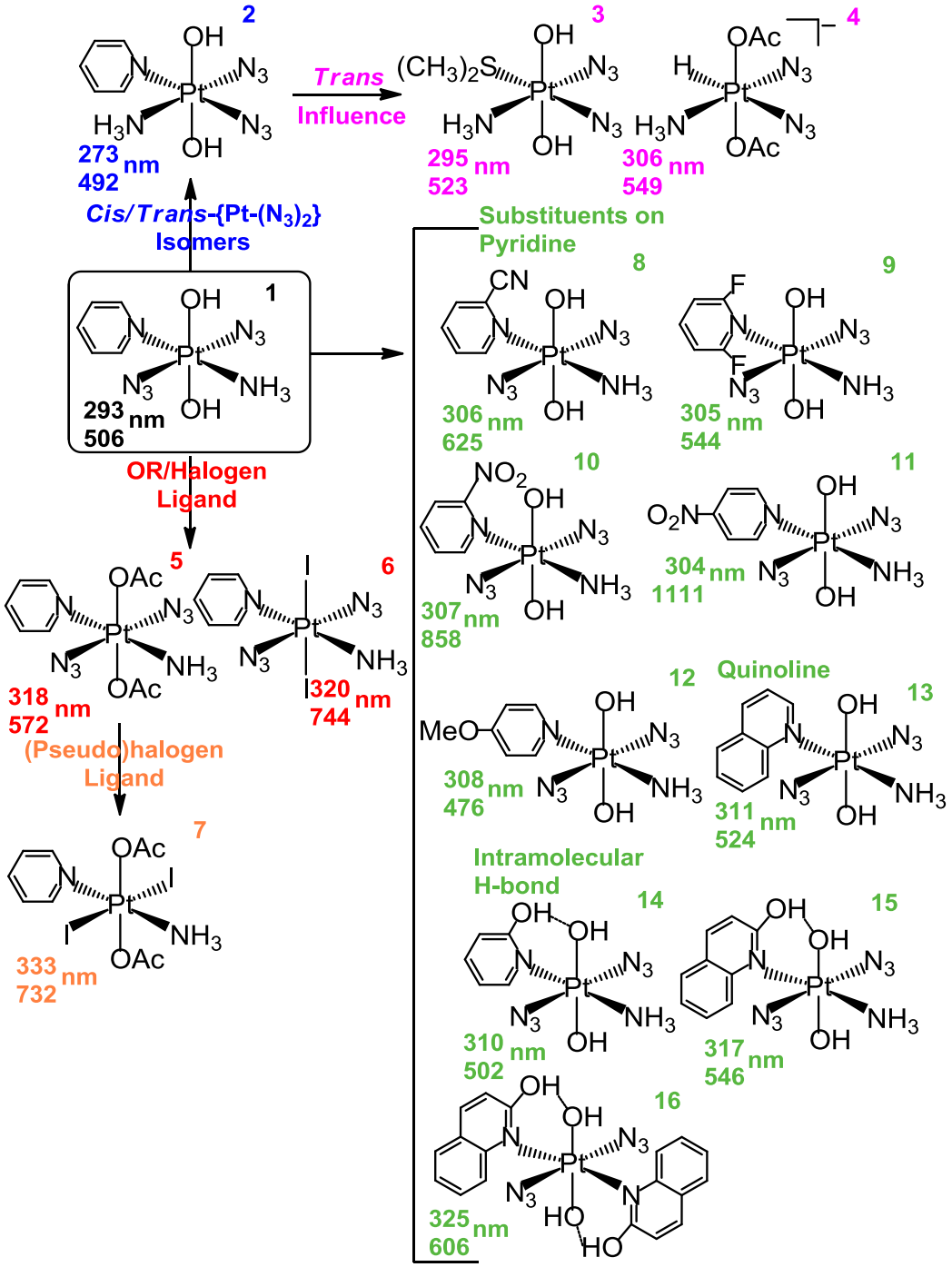


Figure 3-3. Calculated wavelength of the intense band and lowest energy absorption for platinum complexes with various chemical modifications

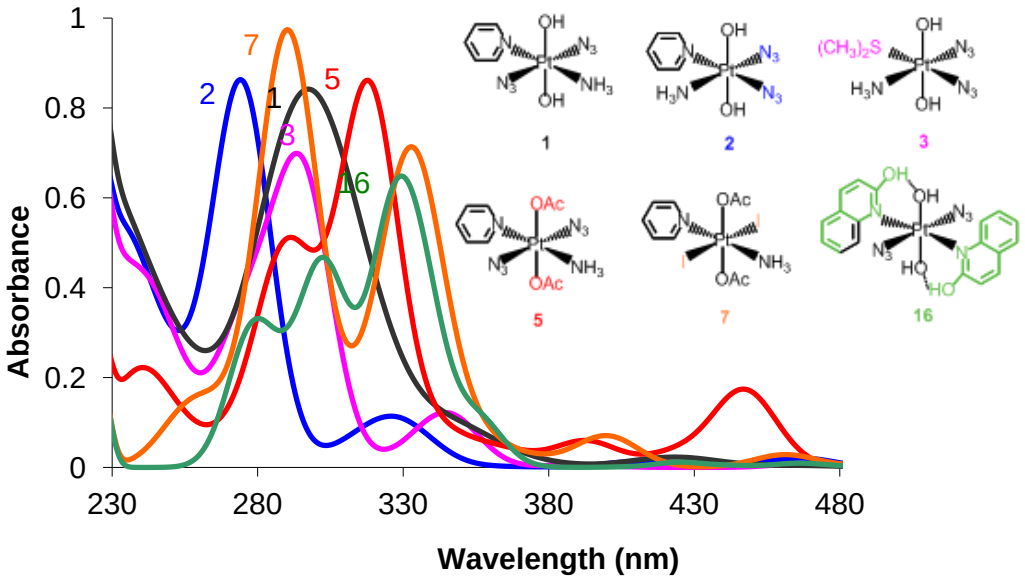


Figure 3-4. Simulated absorption spectra of selected Pt(IV) complexes **1**, **2**, **3**, **5**, **7**, and **16**. The oscillator strength has been scaled by 2.5 times for the convenience of comparing with experimental absorption of complex **1**.

3.3 Orbitals Analyses

In addition to exploring the energies of electronic transitions upon different chemical modifications, it is highly desirable to understand qualitatively how the structural changes in Pt(IV) complexes relate to spectroscopic absorption. By identifying and examining the most important orbitals that control the electronic excitation this correlation and thus mechanisms of photoactivation can be elucidated. The results for electronic transitions of the intense band and electronic transitions of low-energy absorption are summarized in Table 3-2 and Table 3-3, respectively. The major orbital contributions to the main absorption are visualized in Figure 3-5.

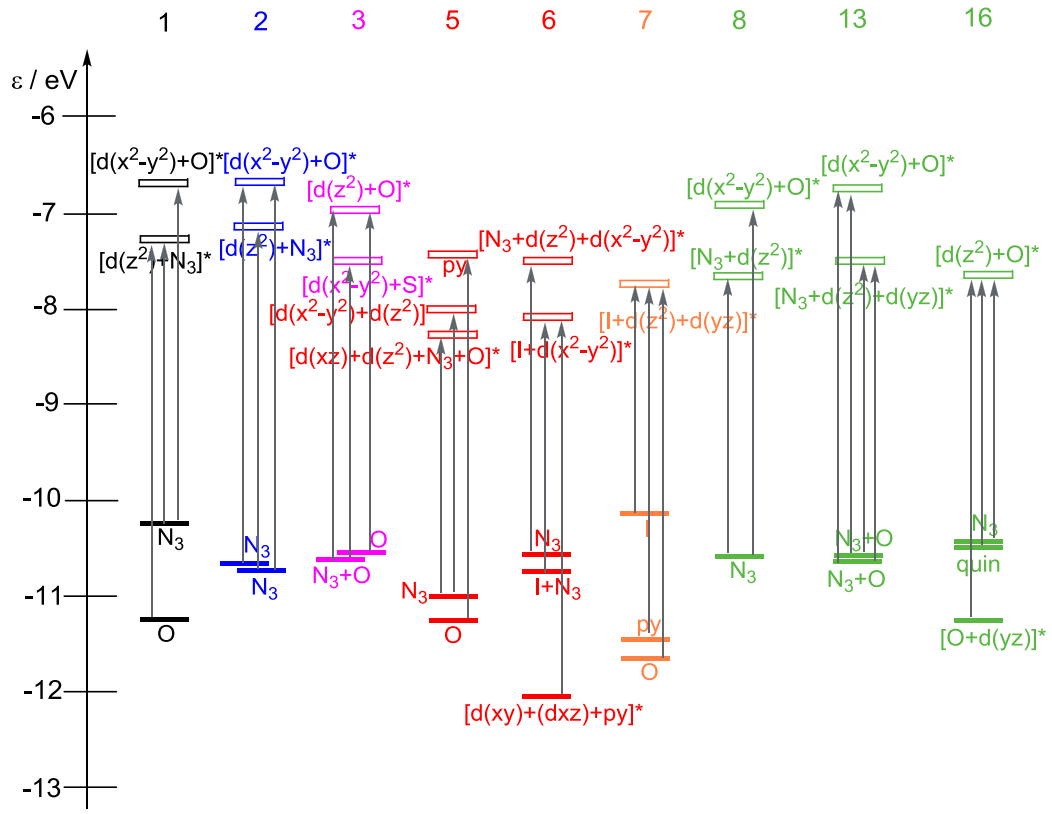


Figure 3-5. Prominent molecular orbital diagram of the main absorption for selected platinum complexes. Orbital energies are given in eV. Full symbols correspond to vacant orbitals; empty symbols correspond to occupied orbitals.

3.3.1 Electronic Transitions of the Intense Band

The key transitions are a mixture of ligand-to-metal charge transfer (LMCT), where electron density migrates from the OR^- and (pseudo)halogen ligands to the metal, and ligand-to-ligand charge transfer (LLCT), where electrons move between OR^- and (pseudo)halogen ligands. In comparison with complex **1** (293 nm), the *cis* isomer, complex **2** (273 nm), has a lower nonbonding orbital of N_3 character resulting in the higher energies of the transitions. The *trans* influence tends to increase the probabilities of electronic transition from N_3 *trans* to Me_2S and slightly lower the energies of the unoccupied orbitals compared to complex **2**.

Table 3-2. Calculated absorption properties (intense band) of Pt(IV) complexes (continued on following pages).

Tr[a]	Composition	Energy/eV (nm)	Oscillator Strength	Assignment
cis, trans, cis-[$Pt(N_3)_2(OH)_2(NH_3)$(pyridine)] (**2**)				
26	HOMO−5→LUMO+2 (47%) HOMO−4→LUMO (18%)	4.50 (276)	0.133	LMCT/LLCT
27	HOMO−4→LUMO+2 (35%) HOMO−5→LUMO (25%) HOMO−5→LUMO+2 (10%)	4.54(273)	0.205	LMCT/LLCT
cis, trans, cis-[$Pt(N_3)_2(OH)_2(NH_3)S(CH_3)_2$] (**3**)				
12	HOMO−4→LUMO+1 (29%) HOMO−4→LUMO (28%) HOMO−3→LUMO+1 (11%)	4.20 (295)	0.255	LMCT/LLCT
13	HOMO−5→LUMO+1 (61%) HOMO−5→LUMO (19%)	4.48 (277)	0.114	LMCT/LLCT
cis, trans, cis-[$PtH(N_3)_2(OAc)_2(NH_3)]^-$ (**4**)				
13	HOMO−2→LUMO+1 (72%)	4.05 (306)	0.103	LMCT/LLCT
17	HOMO−8→LUMO (45%) HOMO−8→LUMO+1 (17%) HOMO−5→LUMO+1 (14%) HOMO−6→LUMO+1 (13%)	4.50 (276)	0.113	LMCT/LLCT
trans, trans, trans-[$Pt(N_3)_2(OAc)_2(NH_3)$(pyridine)] (**5**)				
24	HOMO−4→LUMO (31%) HOMO−4→LUMO+1 (28%) HOMO−5→LUMO+2 (13%)	3.89 (318)	0.243	LMCT/LLCT
	HOMO−7→LUMO (22%) HOMO−7→LUMO+1 (17%)	4.27 (291)	0.148	LMCT/LLCT

	HOMO−8→LUMO+1 (17%)			
trans, trans, trans-[$Pt(N_3)_2I_2(NH_3)$(pyridine)] (**6**)				
31	HOMO−6→LUMO+1 (33%) HOMO−7→LUMO (21%) HOMO−10→LUMO (18%)	3.87 (320)	0.221	LMCT/LLCT/ MLCT/d-d
32	HOMO−6→LUMO+1 (27%) HOMO−7→LUMO+1 (19%) HOMO−10→LUMO (16%)	3.93 (315)	0.292	LMCT/LLCT/ MLCT/d-d
35	HOMO−10→LUMO (48%) HOMO7→LUMO (11%)	4.24 (293)	0.199	LMCT/LLCT/ MLCT/d-d
trans-trans-trans[$Pt(I)_2(OAc)_2(NH_3)$(pyridine)] (**7**)				
25	HOMO−4→LUMO (31%) HOMO−9→LUMO (27%) HOMO−10→LUMO (22%)	3.73 (333)	0.197	LMCT/LLCT
34	HOMO−10→LUMO (43%) HOMO−4→LUMO (21%) HOMO−11→LUMO (13%)	4.28 (290)	0.349	LMCT/LLCT/ MLCT/d-d
trans, trans, trans-[$Pt(N_3)_2(OH)_2(NH_3)$(2-cyanopyridine)] (**8**)				
22	HOMO−3→LUMO+1 (69%) HOMO−3→LUMO+3 (12%)	4.05 (306)	0.333	LMCT/LLCT
trans, trans, trans-[$Pt(N_3)_2(OH)_2(NH_3)$(2,6-difluoropyridine)] (**9**)				
23	HOMO−3→LUMO (43%) HOMO−3→LUMO+2 (33%) HOMO−6→LUMO (15%)	4.07 (305)	0.344	LMCT/LLCT
trans, trans, trans-[$Pt(N_3)_2(OH)_2(NH_3)$(2-nitropyridine)] (**10**)				
28	HOMO−3→LUMO+1 (37%) HOMO−5→LUMO+2 (23%)	4.03 (307)	0.229	LMCT/LLCT/ MLCT

trans, trans, trans-[$Pt(N_3)_2(OH)_2(NH_3)$(4-methoxypyridine)] (**12**)				
17	HOMO−3→LUMO (46%) HOMO−7→LUMO (34%)	4.03 (308)	0.184	LMCT/LLCT
20	HOMO−3→LUMO+3 (64%) HOMO−3→LUMO (19%) HOMO−7→LUMO (12%)	4.26 (291)	0.123	LMCT/LLCT
42	HOMO−5→LUMO+2 (51%) HOMO−9→LUMO+1 (21%)	5.39 (230)	0.168	LMCT/LLCT MLCT/d-d
trans, trans, trans-[$Pt(N_3)_2(OH)_2(NH_3)$(2-hydroxypyridine)] (**14**)				
13	HOMO−3→LUMO (41%) HOMO−4→LUMO (13%)	4.00 (310)	0.253	LMCT/LLCT MLCT/d-d
trans, trans, trans-[$Pt(N_3)_2(OH)_2(NH_3)$(2-hydroxyquinoline)] (**15**)				
17	HOMO−4→LUMO (43%) HOMO−5→LUMO (25%) HOMO−6→LUMO+1 (15%)	3.91 (317)	0.259	LMCT/LLCT MLCT/d-d
38	HOMO−5→LUMO+3 (18%) HOMO−1→LUMO+3 (15%)	5.06 (245)	0.138	LLCT
trans, trans, trans-[$Pt(N_3)_2(OH)_2$(2-hydroxyquinoline)$_2$] (**16**)				
29	HOMO−7→LUMO+1 (35%) HOMO−4→LUMO+1 (21%)	3.81 (325)	0.159	LMCT/LLCT/ MLCT/d-d

[a] Tr = transition number as obtained in the TD-DFT calculation.

The profile of electronic transitions in Figure 5 reveals that electrons in *cis* isomers tend to move from orbitals of the (pseudo)halogen ligands to the d orbitals of the metal, whereas those in all *trans* isomers are more likely to transfer from both OR and (pseudo)halogen ligands to the d orbitals of the metal and thus lead to different photoproducts. The NMR spectra for the photolysis of *trans,*

trans, *trans*-[Pt(N$_3$)$_2$(OH)$_2$(NH$_3$)$_2$] and *cis*, *trans*, *cis*-[Pt(N$_3$)$_2$(OH)$_2$(NH$_3$)$_2$] reported previously correlate with the calculations.[33,34,159] It will be necessary to study a series of *cis*-*trans* isomers and the makeup of the important molecular orbitals in future work for a complete understanding of the mechanisms that give rise to different photopathways and photoproducts.

OAc^- and I^- ligands in Pt(IV) complexes significantly lower the energies of the LUMO and LUMO+1 and lead to a red shift with the main absorption around 320 nm in complexes **5** and **6**. The fairly long wavelength of complex **7** (333 nm) is mainly due to the higher occupied orbital of I^- as leaving ligands yet the energies of unoccupied orbitals are similar to those in complexes **5** and **6**. Otherwise, quinoline and electron-withdrawing substituents on pyridine tend to lower the energies of the unoccupied orbitals of complexes **12**, **8**, **9**, and **10** to a similar level and result in lower energy absorption from 305 to 311 nm, whereas the electron-donating methoxy substituent on aromatic amines reduces the energies of transitions, with a absorption maximum at 308 nm, by increasing the contribution from the frontier orbitals closer to the HOMO and LUMO.

3.3.2 Electronic Transitions of Low-energy Absorption

The TDDFT results in Table 3-3 show that the "tail" of the absorbance extends from the green region for complex **1** (506 nm) to the red region for complexes **8** (625 nm) and **10** (858 nm) which indicates that electron-withdrawal by resonance effect has a larger influence on reducing the energy of the LUMO in Pt complexes compared to electron-withdrawal by the inductive effect and electron-donating substituents on py (544 nm for complex **9**, 476 nm in complex **12**).

Selected frontier orbitals for complex **10** generated by Molekel v. 5.3[166] are presented in Figure 3-5. The main contributions for low-energy absorption are from the nonbonding orbitals for azide ligands (HOMO−2 and HOMO) to the LUMO, which comprises antibonding orbitals on the pyridine nitro group. Some familiar trends are observed in complexes **8** and **16**, which feature unusually low energy absorption. Electron-withdrawing groups tend to increase the likelihood of ligand-to-ligand charge transfer (LLCT), where electrons move from leaving groups to py/quin, leading to a different mechanism of photoreaction. Likewise, the calculated low-energy absorption for complexes **6** and **7** with iodide ligands are 744 nm and 732 nm, respectively. The results are consistent with the UV-Vis spectra of iodide-containing complexes obtained in previous studies.[25] Comparison of the energies of the frontier orbitals in complex **1** reveals that the energy of the LUMO in complex **6** is lowered whereas the energy of the HOMO in complex **7** is raised resulting in a smaller gap between HOMO and LUMO, and thus longer wavelength absorption can be achieved.

Table 3-3. Calculated low-energy absorption properties of Pt(IV) complexes.

Tr[a]	**Composition**	**Energy, eV (nm)**	**Oscillator Strength**	**Assignment**
(1) *trans, trans, trans*-[Pt(N$_3$)$_2$(OH)$_2$(NH$_3$)(pyridine)]				
1	HOMO→LUMO (100%)	2.45 (506)	0.0	LMCT
5	HOMO−2→LUMO (71%) HOMO−1→LUMO+1 (28%)	2.93 (423)	0.007	LMCT/LLCT/ MLCT/d-d
(2) *cis, trans, cis*-[Pt(N$_3$)$_2$(OH)$_2$(NH$_3$)(pyridine)]				
1	HOMO→LUMO (96%)	2.52 (492)	0.0	LMCT
2	HOMO→LUMO+1 (96%)	2.64 (469)	0.007	LLCT

(3) *cis, trans, cis*-[$Pt(N_3)_2(OH)_2(NH_3)(S(CH_3)_2]$				
1	HOMO→LUMO (100%)	2.37 (523)	0.0	LMCT/LLCT
2	HOMO−1→LUMO (50%) HOMO−2→LUMO (48%)	2.65 (468)	0.003	LMCT/LLCT
(5) *trans, trans, trans*-[$Pt(N_3)_2(OAc)_2(NH_3)$(pyridine)]				
1	HOMO→LUMO (96%)	2.17 (572)	0.001	LMCT
7	HOMO−2→LUMO+1 (39%) HOMO−1→LUMO+1 (28%) HOMO−1→LUMO (18%)	2.79 (445)	0.052	LMCT
(6) *trans, trans, trans*-[$Pt(N_3)_2I_2(NH_3)$(pyridine)]				
1	HOMO→LUMO (98%)	1.67 (744)	0.0	LMCT/LLCT
7	HOMO−5→LUMO (54%) HOMO−1→LUMO+1 (37%)	2.20 (564)	0.007	LMCT/LLCT
(7) *trans-trans-trans*[$Pt(I)_2(OAc)_2(NH_3)$(pyridine)]				
1	HOMO→LUMO (99%)	1.69 (732)	0.0	LMCT/LLCT
8	HOMO−1→LUMO+2 (99%)	2.46 (504)	0.002	LLCT
(8) *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(2-cyanopyridine)]				
1	HOMO→LUMO (100%)	1.98 (625)	0.0	LLCT
4	HOMO−2→LUMO (98%)	2.46 (504)	0.003	LLCT
(9) *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(2,6-difluoropyridine)]				
1	HOMO→LUMO (99%)	2.28 (544)	0.001	LMCT
4	HOMO−1→LUMO+1 (98%)	2.45 (506)	0.005	MLCT/LLCT
(10) *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(2-nitropyridine)]				
1	HOMO→LUMO (99%)	1.45 (858)	0.002	LLCT
3	HOMO−2→LUMO (96%)	1.94 (639)	0.014	LLCT
(12) *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(quinoline)]				
1	HOMO→LUMO (97%)	2.37 (524)	0.0	LMCT
6	HOMO−2→LUMO+1 (84%)	2.90 (428)	0.015	LLCT

(14) *trans, trans, trans*-[Pt(N_3)$_2$(OH)$_2$(NH_3)(2-hydroxyquinoline)]				
1	HOMO→LUMO (88%)	2.27 (546)	0.0	LMCT/LLCT
7	HOMO−2→LUMO+1 (98%)	2.88 (431)	0.001	LMCT/LLCT
(15) *trans, trans, trans*-[Pt(N_3)$_2$(OH)$_2$(2-hydroxyquinoline)$_2$]				
1	HOMO→LUMO (98%)	2.05 (606)	0.002	LMCT/LLCT
2	HOMO−1→LUMO (93%)	2.19 (565)	0.007	LMCT/LLCT

[a] Tr = transition number as obtained in the TD-DFT calculation.

3.4 Structures and Bonding Characteristics

The structural features and Pt-L bond lengths of the computed Pt(IV) complexes are listed in Figure A1-1 (Appendix 1) and the results of Pt-N_{py} bond analysis are shown in Table 3-4 (see Section 3.4.2). The relationship between the structures/bonding characteristics and photophysical properties provides insight into the pathway of photoreaction and may help the design of more efficient photoactive anticancer compounds.

3.4.1 Relationship between Calculated Structures and Wavelength of Simulated Intense Absorption Band

Due to the various Pt-binding atoms and the *cis/trans* geometries, the structures of our model Pt(IV) complexes needs to be compared with care (see Figure A1-1, Appendix 1). Nevertheless, in comparing the structural features by groups of the various chemical modifications or by complexes with the same types of atoms coordinated to Pt, the correlation between structures and the absorption maxima, λ_{max} can be revealed. Firstly, for *cis/trans* isomers, the λ_{max} of complex **1** (293 nm) is longer than that of complex **2** (273 nm). The average Pt−N_3 bond distances of 2.09 Å in complex **1** longer (caused by stronger *trans* influence of N_3^- ligands

compared to NH_3 and py) than the corresponding values for complex **2** (2.05 Å) are computed, yet the average Pt−O bond lengths in the two isomers are about 2.05 Å. Secondly**,** the same trend **4** (306 nm; 2.15 Å) > **3** (295 nm; 2.07 Å) > **2** (273 nm; 2.05 Å) is observed for λ_{max} and the average Pt−N_3 bond distances, yet the average Pt−O bond lengths are about 2.07 Å, 2.05 Å, 2.05 Å, respectively. Ligands with a stronger *trans* influence elongate the Pt−L bond for the ligands in the *trans* position, which accounts for the longer Pt-N_3 bond lengths of S−**Pt−N_3** in complex **3** (2.08 Å) and of H−**Pt−N_3** in complex **4** (2.25 Å) compared to that of py-**Pt-N_3** in complex **2** (2.05 Å). On the other hand, ligands with a strong *trans* influence tend to decrease the energy gap between the occupied and unoccupied frontier orbitals in the platinum complexes, leading to a slightly red-shifted intense band in the absorption spectrum (Figure 3-5). In addition, the comparison of the average Pt−$N_{py/quin}$, Pt−N_3, and Pt−O bond lengths as well as λ_{max} values in complex **16** (~2.14, 2.10 and 2.05 Å; 325 nm), complex **5** (~2.08, 2.10, and 2.09 Å; 318 nm), and complex **1** (~2.07, 2.09, and 2.05 Å; 293 nm) shows that the Pt−L bond distances tend to be longer for complexes with absorption maxima at longer wavelength. The relatively shorter Pt−O bond lengths in complex **16** may result from the weaker *trans* influence of OH^- ligands involving intramolecular H-bonds. It is also worth noting that the λ_{max} values of complexes **8–14** are between 304~311 nm and, interestingly, the similar average Pt−N (~2.05**–**2.07 Å), Pt−N_3 (~2.09**–**2.10), and Pt−O (~2.05 Å) in complexes **8–14** are predicted. In general, the comparison of the optimised geometries shows that Pt(IV) complexes with longer Pt−L bonds lengths (particularly for the leaving groups) are more likely to undergo photoreduction with longer-wavelength light.

3.4.2 Energy Decomposition of the Pt–py/quin Bond

To explore the effect of geometry, OR^- and (pseudo)halogen ligands, and various pyridine substituents on the photochemistry of Pt(IV) complexes and on the nature of the Pt-N bond, the Pt-N_{py} and Pt-N_{quin} bond energy can be decomposed into contributions from the strain energy, ΔE_{str} , and the interaction energy, ΔE_{int} ; i.e., $\Delta E = \Delta E_{str} + \Delta E_{int}$. The former is the energy required to deform the equilibrium structures of the metal fragment and py/quin toward their respective geometries in the Pt(IV) complexes. ΔE_{int} is the interaction energy between the in-complex metal and py/quin fragments. It can be further partitioned into three components: (1) the Pauli repulsion between the (non-overlapping) charge clouds of the py/quin and metal fragments at their in-complex positions, ΔE_{Pauli}, (2) the electrostatic interaction between the two fragments, ΔE_{elec}, and (3) the stabilising orbital interactions accompanying relaxation of the electron density, ΔE_{orb}. Thus, $\Delta E_{int} = \Delta E_{Pauli} + \Delta E_{elec} + \Delta E_{orb}$. As illustrated in Table 3-4, relative to complex **1** the largest total interaction energies are observed for complex **14**, with an intramolecular H-bond (**–**45.4 kcal/mol) and complex **12**, with an electron-donating substituent (–40.3 kcal/mol). Electrostatics provides a 30-60 kcal/mol larger stabilizing contribution to the Pt-N_{py} and Pt-N_{quin} bond than do orbital interactions. The ratio of the orbital to the electrostatic contributions, $\Delta E_{orb}/\Delta E_{elec}$, indicates the relative covalent vs ionic character of the Pt-N_{py} and Pt-N_{quin} bond. For Pt(IV) complexes with no intramolecular H-bond, the covalent/ionic ratio, R_{ci}, is between 0.56 and 0.63. The Pt-N_{py} and Pt-N_{quin} bond for complexes with an intramolecular H-bond between OH on the py/quin ring and an OH^- ligand exhibits more covalent bond character (R_{ci} = 0.75–0.80). The strain energy is 6.8 kcal/mol smaller for complex **2** (10.0 kcal/mol) with a *cis* geometry, yet it is

10–20 kcal/mol larger for complex **5** (31.0 kcal/mol) with acetate and complexes **14** (27.6 kcal/mol) and **15** (36.5 kcal/mol) with intramolecular H-bonding compared to complex **1** (16.8 kcal/mol). In terms of bonding stability, these results may make useful suggestions for the synthesis of Pt(IV) complexes.

Table 3-4. Energy decomposition of the Pt-N_{py}[b] and Pt-N_{quin}[c] bond in selected Pt(IV) complexes.

	1	2	5	8	9	12	13	14	15
ΔE_{str}(py)	0.7	1.0	0.4	1.6	1.4	1.2	0.8	9.7	12.0
ΔE_{str}(pt)	16.1	9.0	30.5	20.6	21.3	16.5	20.7	18.0	24.5
ΔE_{str}	**16.8**	**10.0**	**31.0**	**22.2**	**22.7**	**17.7**	**21.5**	**27.6**	**36.5**
ΔE_{Pauli}	167.6	157.0	159.5	162.2	133.9	178.82	166.3	223.8	215.6
ΔE_{elst}	-137.4	-127.6	-131.4	-130.2	-101.4	-148.7	-134.8	-170.1	-158.9
ΔE_{orb}	-84.0	-71.9	-82.0	-79.8	-63.6	-88.1	-83.8	-126.8	-127.4
ΔE_{int}	**-53.7**	**-42.5**	**-53.9**	**-47.7**	**-31.1**	**-58.0**	**-52.3**	**-73.0**	**-70.7**
ΔE_{sum}	**-36.9**	**-32.5**	**-23.0**	**-25.5**	**-8.4**	**-40.3**	**-30.8**	**-45.4**	**-34.2**
$\Delta E_{orb}/\Delta E_{elst}$	0.61	0.56	0.62	0.61	0.63	0.59	0.62	0.75	0.80

[a] Energies in kcal/mol. [b]py = pyridine. [c]quin = quinoline

3.5 Photophysics of *trans, trans, trans*-[Pt$(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] (11)

Tissue-penetration of light is wavelength-dependent. Complexes activated with light of longer wavelength (600–800 nm)[160,167] allow treatment of larger tumours in photochemotherapy. In the calculated UV-Vis spectrum of *trans,trans,trans*-[Pt$(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] (**11**), there is a low-energy low intensity absorption band around 664 nm (Figure 3-7) induced by the electron-withdrawing nitro group, which mainly corresponds to a ligand(azide)-to-ligand(4-nitropyridine) charge-transfer transition. Complex 11 was thus synthesised by Yao Zhao in order to compare its UV-Vis spectrum with the computed results and investigate the photoreaction pathway promoted by the electron-withdrawing substituent.

3.5.1 Experimental Study

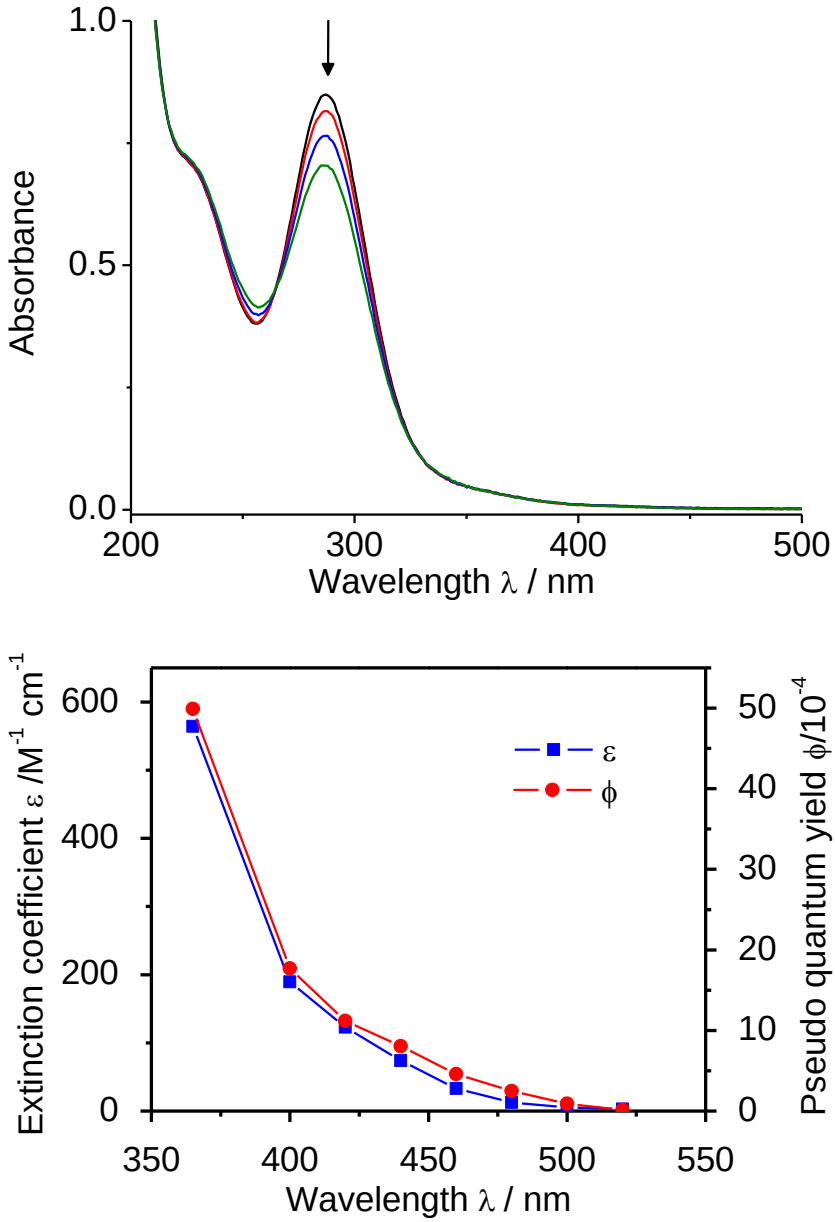


Figure 3-6. (a) UV-Vis spectra recorded for 11 in H_2O upon irradiation of 365 nm (light dose 0, 0.7, 2.1, 4.2 J/cm^2). (b) Pseudo quantum yield ϕ (red, circles) of action spectrum (wavelength-dependent photodecomposition) for 11, in comparison with extinction coefficient ε (blue, squares). The data points are the average of two experiments. Data supplied by Yao Zhao.

Complex **11** *trans,trans,trans*-[Pt$(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] was synthesised and characterised by ^{1}H, ^{13}C, ^{195}Pt NMR and ESI-MS. In order to compare with the calculated spectrum, the UV-Vis spectrum for complex **11** in

aqueous solution was recorded, Figure 3-6a. In the experimental spectrum, there is an absorbance maximum at 287 nm, which is assigned to be a ligand(azide)-to-metal charge-transfer (LMCT) transition. The absorption intensity at >400 nm is very small and no absorption was observed above 540 nm. When the sample was irradiated with a specific wavelength of light (e.g., 365 nm), the major absorption band decreased, indicating the loss of azide groups (Figure 3-6a). For the purpose of investigating the relationship between the wavelength of irradiation and photodecomposition of **11** in aqueous solution as monitored by UV-Vis spectroscopy, an "action spectrum" was measured[36,168] (Figure 3-6b). The action spectrum recorded the rate of photodecomposition of the molecule upon the exposure to light of wavelength 365 to 540 nm. The pseudo quantum yield (ϕ) (ratio of the number of photo reactions to the number of incident photons) was determined (Figure 3-6b). The photolysis of complex **11** (pseudo quantum yield ϕ) upon irradiation from 365 nm to 520 nm correlated closely with the UV-Vis spectrum (extinction coefficient). Over 540 nm, the photolysis was too small to detect with our experimental conditions. The irradiation with light of 600 nm, 633 nm, 660 nm and 700 nm was also carried out individually, but no photodecomposition was observed (data not shown). These experiments were carried out by Yao Zhao.

3.5.2 Computed Spectra and Comparison with Experiment

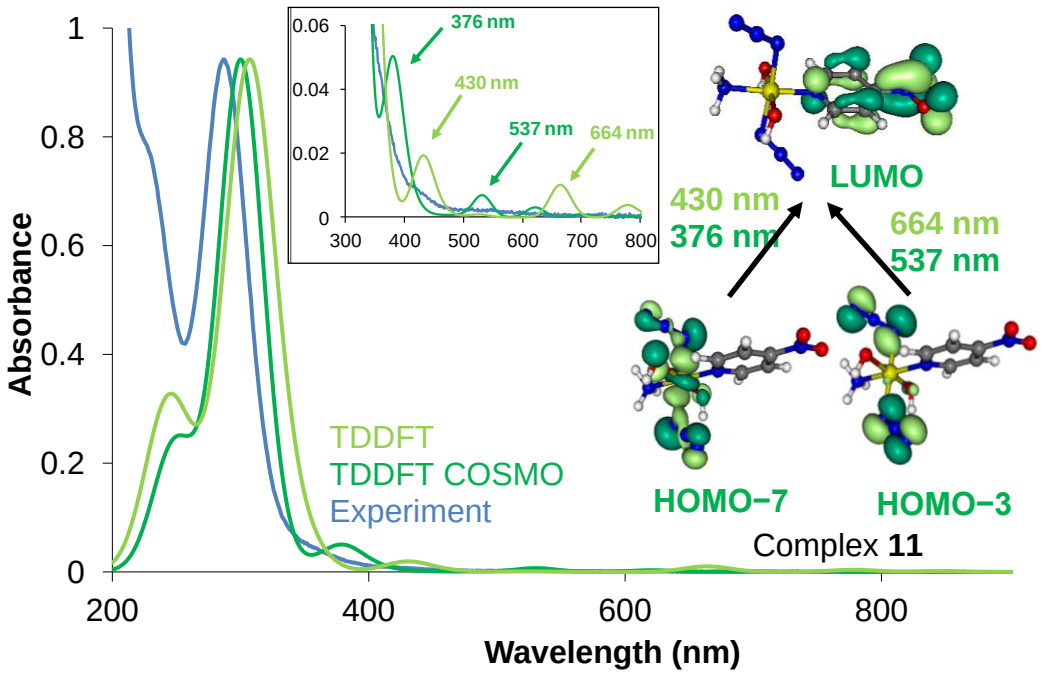
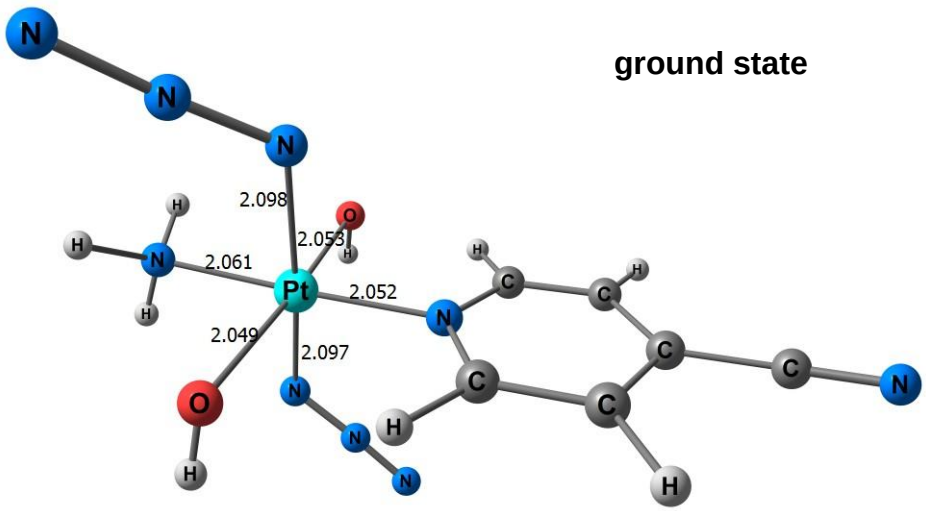


Figure 3-7. Comparison between calculated excited-state transitions and UV absorption for *trans, trans, trans*-[$Pt(N_3)_2(OH)_2(NH_3)$(4-nitropyridine)] (**11**). The oscillator strength has been scaled by 1.84 times for the convenience of comparing with experimental absorption. Representive frontier orbitals are shown on the right. Insert: magnification of the spectra from 300-800 nm. Experimental data supplied by Yao Zhao.

A comparison of calculated and experimental spectra of complex **11** is shown in Table 3-1, Table 3-3, and Figure 3-7. Both spectra show absorption maxima in the UV region, which can be assigned to mixed LLCT/LMCT/MLCT (N_3, OH→N_3, Pt and py, Pt→4-nitropyridine) and LLCT/MLCT (py, Pt→4-nitropyridine) transitions by TDDFT. In the investigation of low-energy absorption, TDDFT predicts two small peaks at 430 nm (LLCT/MLCT; N_3, Pt, OH→4-nitropyridine) and 664 nm (LLCT; N_3→4-nitropyridine), whereas UV-Vis spectrum shows no absorption beyond 540 nm.

The spectrum of complex **11** was therefore recomputed by TDDFT with the Conductor like Screening Model (COSMO)[169] of solvation, and the overlay with the experimental spectrum shows an improvement in terms of intense bands and low-energy absorption. In particular, the two small peaks at 430nm and 664 nm for the low-energy absorption were flattened and two blue-shifted peaks at 376 nm and 531 nm are predicted instead. This increase in energy is mainly due to the stabilization of the occupied frontier orbitals and destabilization of LUMO by solvent effects and the resulting larger energy gap between HOMO−3/HOMO−7 and LUMO (Figure 3-7). This result matches closely the experimental UV-Vis and action spectrum. It is worth to note that the minimum absorption wavelength of complex **11** is *ca*. 20 nm longer in the action spectrum than that of *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ (< 520 nm) in the previous study.[36] This demonstrates that the electron-withdrawing group by resonance effect can extend the absorption to longer wavelength. The photophysical behavior has been shown experimentally to depend on solvent polarity.[170] The measurement of the low-energy low-intensity absorption in less polar solvents or in gas phase may be extended to longer wavelength in closer agreement with the gas phase prediction of TDDFT.



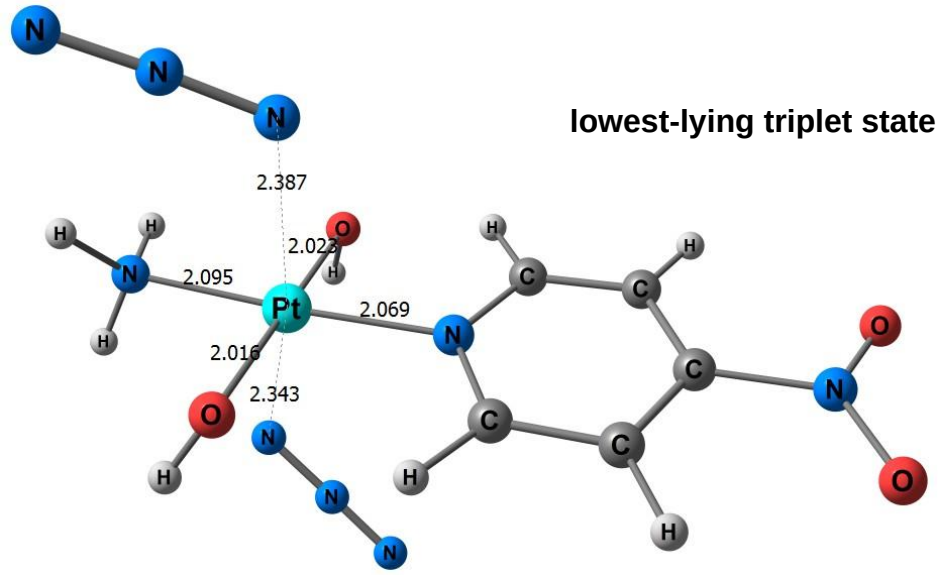


Figure 3-8. The computed geometries of the ground state and the lowest-lying triplet state for *trans, trans, trans*-[Pt(N_3)$_2$(OH)$_2$(NH_3)(4-nitropyridine)] (**11**).

In addition, the lowest-lying triplet states of transition metal complexes have been shown experimentally to make prominent contributions to the photodissociation process.[171] The geometry of the lowest-lying triplet state of complex **11** was thus optimised and is displayed in Figure 3-8. Relative to the singlet ground state, the two Pt−N_3 bond lengths are significantly elongated from 2.098 Å and 2.097 Å to 2.387 Å and 2.343 Å in the triplet state (Figure 3-8). Mulliken population analysis[172] shows that the two nitrogen atoms of azide coordinated to Pt have spin density of 0.23 and 0.28 e. The results of the calculation correlate with the release of azido ligands as N_3• radicals observed experimentally.[159]

3.6 Conclusions

In this work, the photochemistry of Pt(IV) complexes has been investigated in terms of the coordination environment in relation to electronic excitation energies by DFT and TDDFT. The main findings and the biological relevance of this work

are as follows: (1) Calculated excitation energies for the intense band according to TDDFT with the SAOP functional are in good agreement with experimental observations; the theory has been shown to be a useful tool to portray spectroscopic absorption, to further assist the design of Pt(IV) complexes to achieve photoactivation at longer wavelength, and to study the mechanisms of photoreaction (*vide infra*). (2) Electronic transitions for *cis*-{Pt(IV)$(N_3)_2$} isomers tend to arise from the orbitals of the (pseudo)halogen ligands to the d orbitals of metal, whereas those for all *trans* isomers are more likely to arise from both OR and (pseudo)halogen ligands to the d orbitals of the metal and thus lead to different photoproducts. (3) Electron-withdrawing/donating substituents on pyridine tend to slightly red shift the wavelength of the main absorption compared to the unsubstituted (e.g. complex **1**) by up to 15 nm. However, electron-donating substituents on pyridine reduce the energies of transitions by increasing the contribution from frontier orbitals closer to the HOMO and LUMO while quinoline and electron-withdrawing substituents on pyridine achieve absorption at longer wavelength by lowering the energies of the unoccupied orbitals. (4) The design of photoactivated Pt drugs with slower reduction rates and greater absorption at longer wavelengths can be better achieved via lowering the energies of the unoccupied frontier orbital of Pt complexes. Intramolecular hydrogen bonding involving the OH ligand and 2-hydroxyquinoline red shifts the main absorption by 24–32 nm and extends the lowest energy absorption toward the red region by 40–100 nm. Moreover, mesomeric electron-withdrawing groups (CN^- and NO_2^-) on pyridine also show low-energy absorption with significant intensity in the visible region. (5) The Pt(IV) complexes with longer calculated Pt−L bond lengths tend to absorb light of longer wavelengths. (6) A comparison of the

computed and experimental UV-Vis spectra of complex **11** shows that the prediction of low-energy absorption in solution can be improved by using a solvent model, whereas the absorption maximum is little affected. The absorption of the complexes in clinical treatment may lie between the results of the TDDFT with and without a solvent model depending on the solvent-accessibility in the tissues. (7) The structural and spin density analysis of the lowest-lying triplet state for complex **11** suggests the release of azido ligands as N_3• radicals is possible, which is consistent with the experimental data.

Chapter 4

Ligand Field Molecular Mechanics Modelling of Platinum–DNA Interactions

4.1 Introduction

Computational approaches can provide insights that are not fully accessible experimentally. However, quantum mechanics (QM) is too expensive for large systems like proteins and DNA which contain significant numbers of atoms. Hence, classical techniques such as molecular mechanics (MM) and molecular dynamics (MD) are usually used for macro-biomolecules.[70,113] MD simulations can generate a trajectory that provides atomic details of the time evolution of the system. MD is based on a simple function of nuclear positions via empirical force fields and has a successful record in the study of proteins and nucleic acids. However, MD simulations of metal-containing systems present significant challenges since molecular mechanics ignores electronic motions, yet d electrons can have a potent effect on the structure and properties of coordination complexes. In response, QM/MM is being increasingly used for biomolecules with metal centers, although the computational cost of the QM part remains a major bottleneck. Alternatively, the flexibility and precision of quantum mechanics at the speed of molecular mechanics can be achieved by ligand field molecular mechanics (LFMM) which has been applied to the modeling of electronically 'difficult' metal ions such as d^9 Cu(II) and both spin states of d^8 Ni(II) centres for which conventional MM does not provide a general approach.[72,98,100]

Conventional force fields have been applied to Pt-guanine binding using torsional constraints or out-of-plane deformation terms to ensure the Pt atom remains within the plane of the coordinated purine. However, *ab initio* calculations[71,173] suggest the platinum out-of-plane bending force constant is overestimated. Here, I first develop new LFMM parameters for guanine bound to platinum via N7,

which, in combination with previous work on modeling of the *trans* influence in Pt(II) complexes,[173] reproduces experimental bond lengths and bond angles of small platinum complexes.

4.2 Results of Pt-guanine LFMM Parameterisation

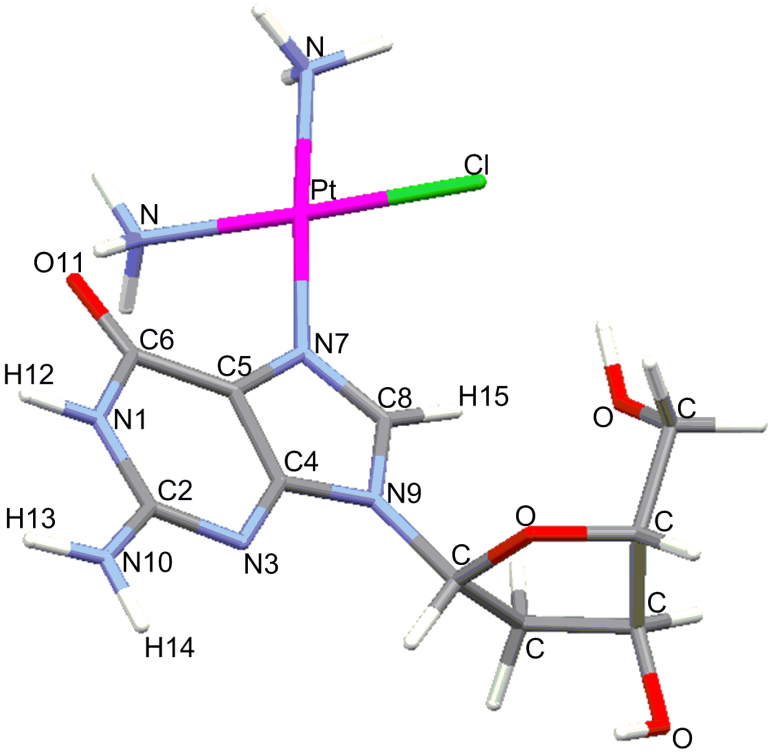


Figure 4-1. Model compound, *cis*-$[Pt(NH_3)_2(Cl)(dGuo)]^+$, used to derive the new charge scheme for Pt-binding guanines.

To determine the angular coordination of L–Pt–L interactions and model intermolecular interactions, partial charges for coordinated guanines are required. These were derived by adding the change in CHelpG partial atomic charges between isolated and Pt-coordinated guanine to the standard MMFF94[106-110] or AMBER94[91] FF partial charges for uncoordinated guanines (computational details in Section 2.5.2). Figure 4-1 displays the complex, *cis*-$[Pt(NH_3)_2(Cl)(dGuo)]^+$,

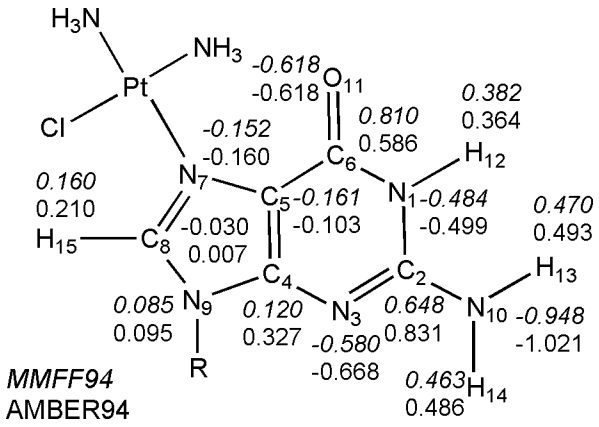


Figure 4-2. New LFMM/MMFF94 and LFMM/AMBER94 charge schemes for N7-platinated deoxyguanosine (dGuo(Pt)$^{+}$), derived from CHelpG charges of the model compound, *cis*-[Pt(NH_3)$_2$(Cl)(dGuo)]$^{+}$, in Figure 4-1.

used for developing the new LFMM/MMFF94 and LFMM/AMBER94 charge scheme (see Figure 4-2 and Table A2-1 in Appendix 2). There charges were employed for coordinated guanines in this work. The charge on Pt, q_{Pt}, can be computed from the total charge transferred by all the ligands to Pt(II) , ΔQ:

$$q_{Pt} = 2 - \Delta Q \qquad (4\text{-}1)$$

Having fixed the partial charge scheme, LFMM requires Morse function parameters for Pt–L bond stretching (r_0, α, D_0), and ligand–ligand repulsion parameters (A_{LL}) for L–Pt–L angle bending. The Cambridge Structure Database[174] (CSD) was searched for X-ray structures of complexes that could be used to optimize new Pt-guanine parameters. Fifteen complexes were selected and, in conjunction with the existing LFMM parameters for Pt-NH_3 and Pt-Cl,[173] Pt-guanine parameters were optimized to reproduce experimentally observed structures with a target deviation in all Pt–L bond lengths of less than 0.05 Å and G–Pt–G bond angles of less than 3° (Table 4-1). The Morse function is of the following form:

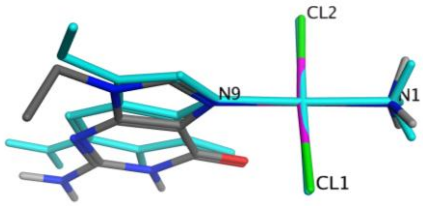
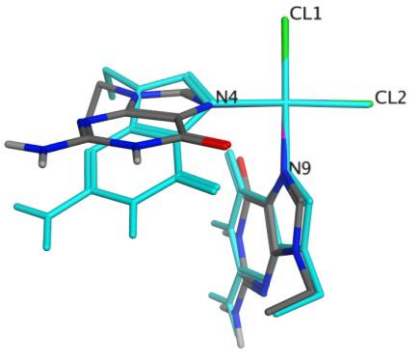
$$\mathrm{E_b(Morse)} = D_0[1 - \mathrm{e}^{-\alpha(r-r_0)}]^2 \quad (4\text{-}2)$$

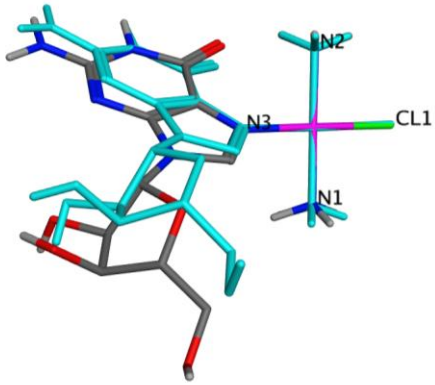
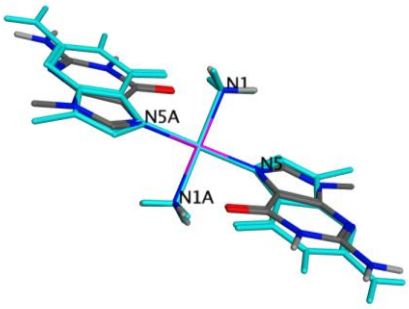
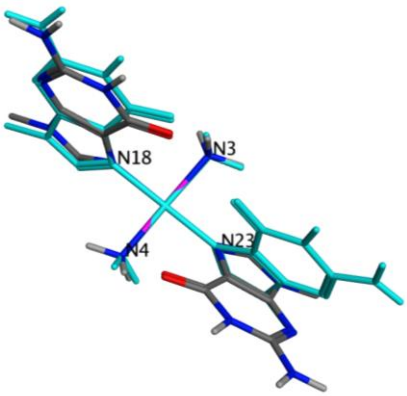
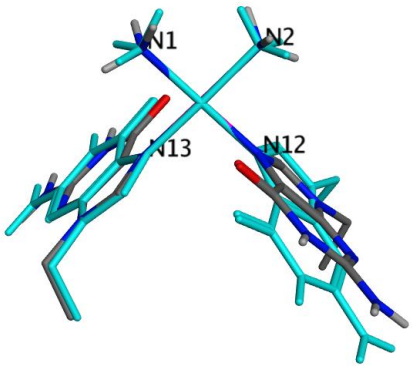
where r_0= 2.241 Å, α= 1.728 Å^{-1}, D_0= 60 kcal/mol were obtained in the parameterization. The LFMM parameters for platinum complexes are available in Table A2-2 in Appendix 2. Apart from *cis*-diammine-bis(9-ethylguanine)-platinum(II) (DEGXA0) where LFMM places one of the ammine ligands 0.075 Å further from the metal than observed experimentally, the only apparent major discrepancy is with diammine-(deoxy-guanylyl-guanylyl-N^7,$N^{7'}$)-platinum(II) (DETLIX). Two of the five CSD entries are included in Table 4-1, DETLIX01[175] and DETLIX20[176], which shows fairly large variations of what should be the same bonds. This leads to the conclusion that the differences between LFMM and experiment are largely attributable to the quality of the experimental data. Several disorder issues are reported in the other structures. Further justification for our assertion comes by comparing LFMM results with previously reported QM/MM data.[74] The LFMM Pt–L bond lengths and G–Pt–G bond angles for square-planar platinum complexes and DNA fragments accurately reproduce QM/MM (Table A2-3, Appendix 2), while overlays of selected molecules confirm that the agreement extends to entire systems (see Table 4-1). The largest absolute difference between LFMM and QM/MM Pt-L contacts is 0.03 Å with the average error ~0.02 Å. In addition to structural features, LFMM parameters have been used for a vibrational analysis of *cis*-$[Pt(NH_3)_2(Cl)(dGuo)]^+$ (Figure 1) which was compared to the BP86/TZ2P results. The computed values of the low energy LFMM frequencies are comparable with the respective quantum mechanical values (Table A2-4, Appendix 2). While the detailed atomic motions vary somewhat, the similarity of the energies suggests that the LFMM and DFT force

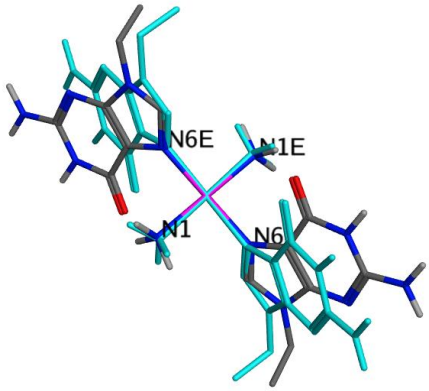
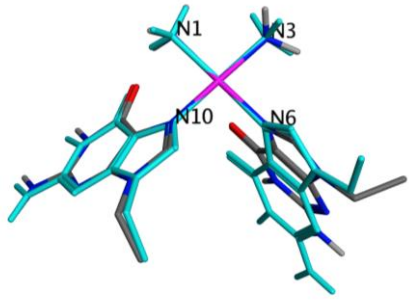
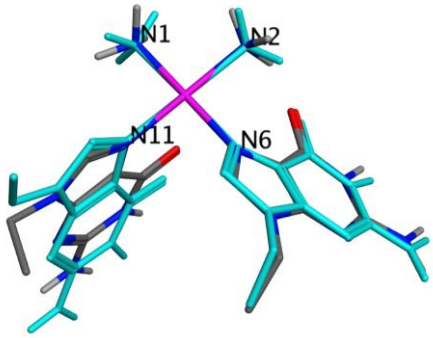
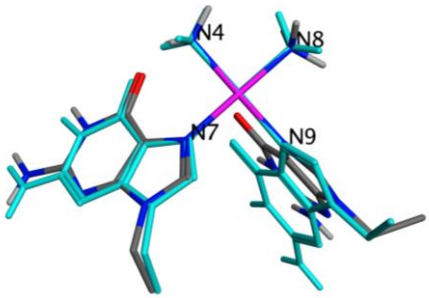
constants are comparable and that the subsequent LFMD trajectories should be realistic.

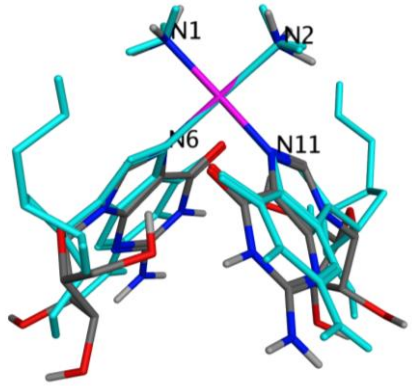
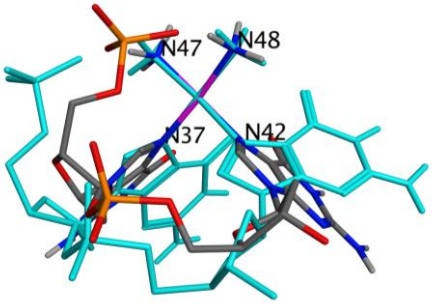
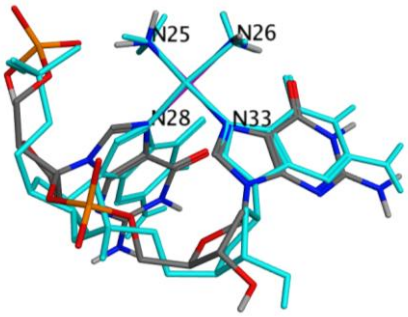
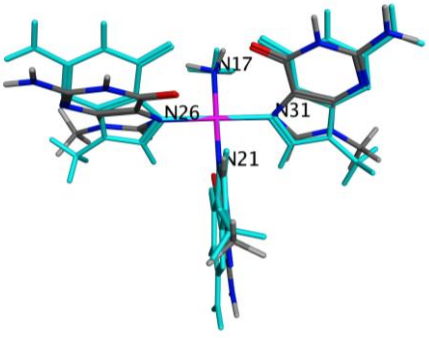
This initial parameter development employed the MMFF94 force field since it has a richer set of ligand types. However, for DNA simulations, I wish to use the AMBER force field. Fortunately, transferring LFMM parameters from MMFF94 to AMBER is straightforward, as shown previously and involves only some minor adjustments of the partial charges and renaming of the atom types.[77]

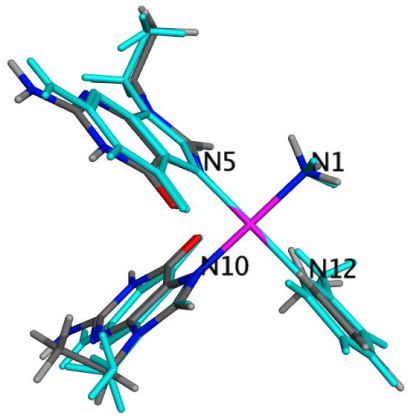
Table 4-1. LFMM-Optimized (color by element; Pt in magenta) and Experimental (cyan) Pt–L Bond Lengths (in Å) for Pt^{2+} Complexes containing Guanine

.Structure (CCDC ref no)		LFMM	Experiment	Δ(LFMM − EXP)
JAQCUZ 	N1 N9 Cl1 Cl2	2.022 1.985 2.304 2.304	2.025 2.007 2.306 2.291	-0.003 -0.022 -0.002 0.013
SEWDUV 	N4 N9 Cl1 Cl2	2.021 2.021 2.317 2.317	2.026 2.005 2.292 2.269	-0.005 0.016 0.025 0.048

HEKJAI 	N3 N1 N2 Cl1	2.037 2.028 2.024 2.272	2.001 2.037 2.050 2.294	0.036 -0.009 -0.026 -0.022
TUWTAG 	N5 N5A N1 N1A	2.011 2.011 2.037 2.037	1.993 1.993 2.035 2.035	0.018 0.018 0.002 0.002
TUWTEK 	N18 N23 N3 N4	2.014 2.014 2.038 2.038	2.041 1.989 2.052 2.030	-0.027 0.025 -0.014 0.008
ZODGII 	N12 N13 N1 N2	2.001 2.001 2.042 2.042	2.011 1.985 2.059 2.034	-0.010 0.016 -0.017 0.008

JUJFID 	N6 N6E N1 N1E	2.011 2.011 2.037 2.037	2.019 2.020 2.033 2.032	-0.008 -0.009 0.004 0.005
CUHJOE 	N6 N10 N1 N3	2.003 2.013 2.039 2.044	2.017 2.025 2.035 2.052	-0.014 -0.012 0.004 -0.008
DEGXA0 	N6 N11 N1 N2	2.013 2.003 2.044 2.039	1.963 2.010 1.969 2.046	0.050 -0.007 0.075 -0.007
CUHJIY 	N7 N9 N4 N8	2.013 2.003 2.039 2.044	2.023 2.029 2.046 2.040	-0.010 -0.026 -0.007 0.004

GUOAPT20 	N6 N11 N1 N2	2.011 2.009 2.045 2.046	1.993 2.018 2.088 2.070	0.018 -0.009 -0.043 -0.024
DETLIX01 	N37 N42 N47 N48	2.008 1.996 2.031 2.037	1.936 2.090 2.086 2.035	0.072 -0.094 -0.055 0.002
DETLIX20 	N28 N33 N25 N26	1.992 2.030 2.070 2.043	2.020 2.054 2.008 2.061	-0.028 -0.024 0.062 -0.018
LOSYOH 	N21 N26 N31 N17	1.994 2.007 2.012 2.034	2.030 2.011 2.004 2.045	-0.036 -0.004 0.008 -0.011

KARYEI 	N5 N10 N1 N12	2.010 2.004 2.049 2.030	2.008 2.017 2.041 2.020	0.002 -0.013 0.008 0.010
	Rmsd (Å)			0.028

4.3 Discussion

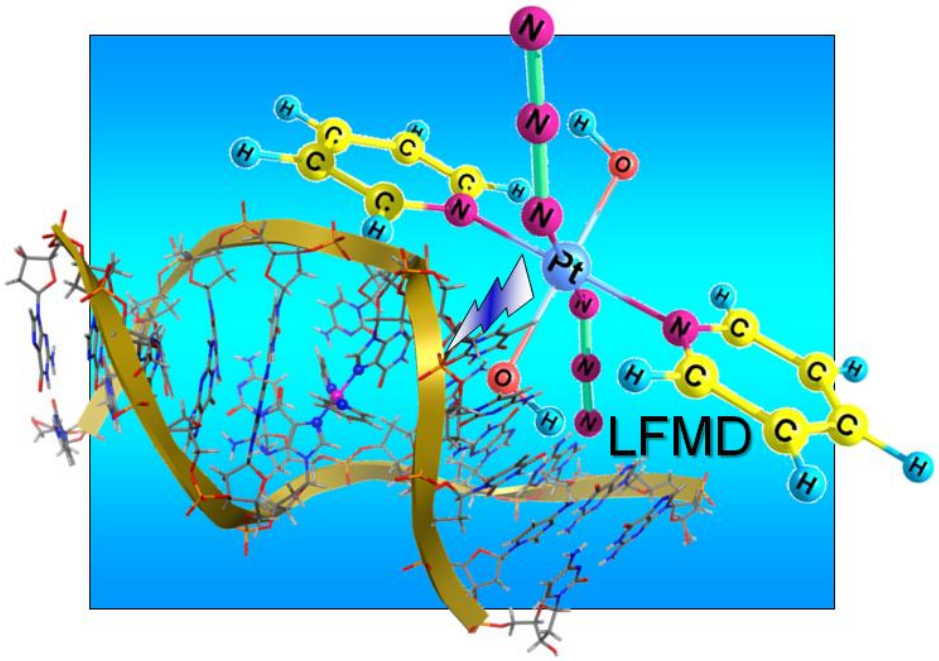
In this work, novel LFMM parameters for DNA simulations of Pt(II) bound to the N7 position of guanines have been developed. The LFMM approach differs from conventional MM by explicitly taking into account the ligand field stabilization energy (LFSE). By parameterizing the ligand field in terms of the bond-centered angular overlap model, a sophisticated description of metal-ligand bonding is possible which captures most of the essential physics around the metal centre. For low-spin d^8 Pt(II) systems such as those studied here, the square-planar geometry of the platinum fragment is generated automatically rather than imposed by artificial forces. Also, the computational efficiency of LFMM allows relatively extensive molecular dynamics simulations from which a detailed description of global DNA parameters can be extracted. Compared to QM/MM, LFMD appears to afford closely comparable accuracy but at a small fraction of the computational cost.

4.3.1 Accuracy of LFMM

The new LFMM parameters are able to reproduce crystal and QM/MM structures in terms of Pt–L bond lengths and Gua–Pt–Gua bond angles for square-planar platinum–DNA complexes, except for one system (CSD code DETLIX01: polymorph I and DETLIX20: polymorph II, see Table 4-1). The existence of different crystal forms is presumably a result of differences in the conditions of crystallization. However, the relatively large difference of over 0.16 Å in Pt–N(guanine) bond lengths in DETLIX01 is clearly at odds with the rest of the data in Table 4-1, which suggest a fairly narrow range of Pt-N7 distances. The difference might be due to the tightly-packed aggregate comprised of four crystallographically independent molecules held together by an extensive network of hydrogen bonding and intermolecular base-base stacking interactions reported for the structure of DETLIX01. However, on average, the experimental Pt–NH_3 and Pt–N7 bond lengths are 2.04 Å and 2.00 Å, respectively,[175] which are in good agreement with the corresponding LFMM values and a difference of 0.16 Å in the Pt-N distances seems anomalously large. I conclude that there is probably an error in the DETLIX01 sturcture.

Chapter 5

Ligand Field Molecular Dynamics Simulations of Platinum–DNA Cross-links: a Distinctive *Trans* DNA Conformation Formed by a Photoactivated Platinum Complex



5.1 Introduction

The new LFMM parameters developed in Chapter 4 should provide a reliable basis for modelling the highly strained structures expected for Pt-DNA binding. In this Chapter, they are used in simulations of (i) 1,2-intrastrand cisplatin–DNA adducts in which *cis*-$\{Pt(NH_3)_2\}^{2+}$ is coordinated to the N7 atoms of adjacent guanine bases, (ii) interstrand cisplatin–DNA adducts in which *cis*-$\{Pt(NH_3)_2\}^{2+}$ is bound to the N7 position of guanines on both strands of duplex, and (iii) a **P**–DNA complex, comprising a *trans*-$[Pt^{II}(py)_2(dGua)_2]^{2+}$ interstrand cross-link formed by binding of the major photoreduction product of the novel platinum(IV) complex *trans,trans,trans*-$[Pt^{IV}(N_3)_2(OH)_2(py)_2]$ (**17**) to two nucleobases in DNA. Among various platinum–DNA products, 1,2-intrastrand adducts are the major DNA products formed by cisplatin and are thought to be important for its anticancer activity, while the minor interstrand cross-links prevent DNA strand separation and can block DNA replication and/or DNA transcription and can be lethal to cells lacking the ability to remove the cross-link.[177] These systems have also been structurally characterized and serve to validate our ligand field molecular dynamics (LFMD) protocols.

The Pt(IV) complex **17** is stable in the dark, but has remarkable cytotoxicity when activated with a low-dose (5 J cm^{-2}) of visible light (420 nm). In the presence of 5'-GMP, the major photoproduct is *trans*-$[Pt^{II}(py)_2(5'\text{-}GMP)_2]^{2+}$ as confirmed by ^{1}H and ^{195}Pt NMR spectroscopy.[36] Photoactivation promotes the rapid formation of such *bis*-guanine adducts whereas in thermal reactions involving *trans* diamine complexes such *trans* cross-links are formed very slowly, if at all.[178] The unusual nature of these *trans*-GG crosslinks formed on DNA after photoactivation of

trans-diazido Pt^{IV} complexes may contribute to their high potency towards cancer cells.[35]

The detailed molecular mechanisms by which cells process interstrand cross-links are not understood. Recently, experiments have shown that formation of 1,2-d(GpG), and 1,3-d(GpTpG) intrastrand cisplatin–DNA adducts in the nucleosome core can alter nucleosome mobility[179] and/or positioning[180,181] and compounds with higher molecular weights display a greater tendency to target the linker region of nucleosomal DNA.[182] Brabec et al have now demonstrated experimentally that the predicted *trans*-$[Pt^{II}(py)_2(G)_2]^{2+}$ interstand crosslink is formed in significant amounts when complex **17** is photoactivated in the presence of DNA. Our model for the major platinum–DNA product formed by photoactivation of complex **17** is thus based on experimental observations in an attempt to contribute to an understanding of the mechanism of action of this novel platinum lesion.

5.2 Calculations

As AMBER has been extensively applied to simulations of biomolecules and the derived partial charges in AMBER94[91] and MMFF94[106-110] are in close agreement (*vide infra*), DommiMOE LFMD simulations were carried out at physiological pH and a mean temperature of 300 K using the LFMM/AMBER94 force field on (i) a 1,2-intrastrand cisplatin–DNA dodecamer containing 767 atoms solvated in a droplet of 1782 water molecules (~10 Å thick) and (ii) an interstrand cisplatin–DNA decamer containing 640 atoms solvated in a droplet of 1531 water molecules. To explore the influence of periodic boundary conditions,[113] the LFMM method was ported to DL_POLY2[183] and the simulation of (iii) the

interstrand cisplatin–DNA complex containing 640 atoms solvated now with 3941 water molecules in a periodic box of dimensions 48x45x56 Å , and, finally, (iv) an interstrand **P**–DNA dodecamer formed by the major photoproduct of complex **17** with DNA, which contains 781 atoms solvated with 3400 water molecules in a periodic box of dimensions 45x45x60 Å .

5.2.1 LFMD Protocol

The starting point for the 1,2-intrastrand cisplatin–DNA adduct simulations was the solution NMR structure (determined by restrained molecular dynamics refinement) of the DNA duplex dodecamer (PDB entry 1A84)[10] d($C_1C_2T_3C_4T_5G_6$*G_7*$T_8C_9T_{10}C_{11}C_{12}$)•d($G_{13}G_{14}A_{15}G_{16}A_{17}C_{18}C_{19}A_{20}G_{21}A_{22}G_{23}G_{24}$), containing a *cis*-GG adduct (where G* denotes the location of platinated nucleotides). The starting point for the interstrand adduct simulations was the 1.63 Å resolution X-ray structure of the double-stranded DNA decamer (PDB entry 1A2E),[9] d($C_1C_2T_3C_4G_5$*$C_6T_7C_8T_9C_{10}$)•d($G_{11}A_{12}G_{13}A_{14}G_{15}$*$C_{16}G_{17}A_{18}G1_9G_{20}$), containing a cisplatin interstrand cross-link. The Protonate3D application in MOE was used to assign ionization states and add hydrogens (as described in Section 2.4.3.1). The structure was immersed in a 10–Å –thick layer of TIP3P water molecules of density ~1 g/cm^3 along with sodium ions used to balance the negative charge of each phosphate. The cisplatin/DNA system was frozen and the solvent molecule positions were energy-optimized. With cisplatin/DNA still fixed, the system was then heated to 300 K in two rounds of 10 ps by coupling to a heat bath using the Berendsen algorithm followed by 100 ps simulation to equilibrate the water at a temperature of 300 K maintained by a Nosé–Hoover thermostat. All atoms were then optimized to give the starting point for LFMD simulations followed by 5 ns of production dynamics under NVT condition. All the atoms

were propagated according to Newton's equations of motions with a time step of 2 $\times 10^{-15}$ s at a mean temperature of 300 K using a Nosé–Hoover thermostat. Bonds involving hydrogen atoms were constrained during the simulations using the SHAKE algorithm.[118] Nonbonded cutoffs (r_1) of 10 Å and an onset (r_0) of 8 Å (the MOE default) were applied to the nonbonded energy terms with the smoothing function:

$$s(r) = \begin{cases} 1 & if\ r < r_0 \\ 1 - [x^3(6x^2 - 15x + 10)] & if\ r_0 \le r \le r_1 \\ 0 & if\ r > r_1 \end{cases}, \qquad (2)$$

where $x = (r - r_0)/(r_1 - r_0)$.

The starting structure for **P**–DNA complex simulations was generated from the 2.16 Å resolution X-ray structure of the DNA duplex dodecamer (PDB 3CO3,[184] d($C_1C_2T_3C_4T_5C_6G_7$*$T_8C_9T_{10}C_{11}C_{12}$)•d($G_{13}G_{14}A_{15}G_{16}A_{17}C_{18}G_{19}A_{20}G_{21}A_{22}G_{23}G_{24}$). This contains a monofunctional adduct of *cis*-$\{Pt(NH_3)_2(py)\}^{2+}$ with a central G after replacement of the original metal fragment with the P fragment, *trans*-$\{Pt(py)_2\}^{2+}$, and formation of the interstrand *trans*-$[Pt(py)_2(dGua)_2]^{2+}$ cross-link between the N7 positions of the central guanines on the opposite strands of DNA, the metal and its coordinated groups were optimized with the rest of the residues frozen. The resulting structure was then reoptimized via multiple rounds of simulated annealing at 450 K to give the starting point for NPT simulations followed by 6 ns of production with a time step of 2 fs at 300 K and a pressure of 1 atm maintained by Berendsen baro- and thermostats using DL_POLY. Long range interactions were treated by using the Ewald summation method. The SHAKE algorithm was employed to constrain all bonds involving hydrogen atoms in water molecules. A cutoff of 12 Å was applied to the nonbonded interaction.

The effect of counterions on DNA structural and thermodynamic properties has been implicated in diverse experiments. A recent study showed that although simulations longer than 60 ns are necessary for ion convergence, the internal structural parameters in MD on DNA are well converged at ≈5 ns.[97] DommiMOE was used to generate input files for DL_POLY simulations.

5.3 Results

5.3.1 Droplet LFMD Simulation of 1,2-intrastrand Cisplatin–DNA Complexes

Table 5-1. Comparison between Selected LFMD, QM/MM,[75] and Experimental Pt Bond Distances and Angles and atomic RMSD for 1,2-intrastrand Cisplatin–DNA Complex PDB 1A84

method	N7–Pt–N7 (deg)	Pt–N (Å)	Pt–N' (Å)	Pt–N7 (Å)	Pt–N7' (Å)	RMSDs[a] (Å)
NMR	90.1	2.049	2.050	2.046	2.051	
QM/MM	86	2.11	2.09	2.06	2.03	0.04/1.50/4.20
DommiMOE	88.9	2.022	2.024	1.986	2.011	0.041/0.554/0.917

[a]RMSDs for Pt–L bond lengths, heavy atoms of $\{Pt(NH_3)_2(Gua)_2\}$ and heavy atoms of whole cisplatin–DNA complex

The DommiMOE simulation of the 1,2–intrastrand cisplatin–DNA adduct started from the NMR conformation (PDB code 1A84) to which counterions and water molecules were added as described above. The potential energy converges after ~2500 ps (within 0.2%, Figure A3-1, Appendix 3). The RMSD of the heavy atoms from the starting structure initially rose, then plateaued after ~2500 ps, and fluctuated around a mean of 3.05 ± 0.11 Å in the simulation. Thus, the first 2500 ps of the trajectory were excluded, while the remaining 2500 ps were used in computing the average structure, distances, and angles. A comparison between LFMD, previous QM/MM results[75] and experimental data is given in Table 5-1.

These include three root-mean-square deviations (RMSDs) for, respectively, Pt–L bond lengths, the heavy (non-hydrogen) atoms of Pt(II) and its ligands NH_3 and Gua, and of the entire Pt-DNA system. Figure 5-1 displays an overlay of LFMM and experimental structures. Overall, all simulations yield average N7–Pt–N7 angles and Pt–N distances that are close to those found for the cisplatin binding site in the experimental structure. The DommiMOE LFMD RMSD for Pt–L bond lengths and heavy atoms of cisplatin–guanine and of cisplatin–DNA complexes are 0.041, 0.554, and 0.917 Å, respectively. The QM/MM simulations for the same adduct have RMSDs of 0.04, 1.50 and 4.20 Å.[75]

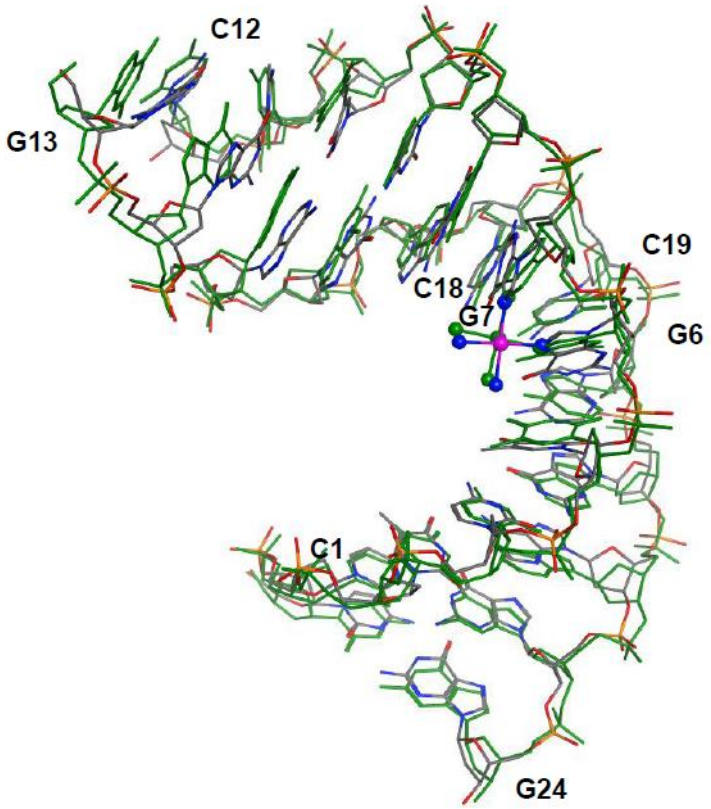


Figure 5-1. NMR structure (green) of the 1,2-intrastrand Cisplatin–DNA Complex: PDB 1A84 with the sequence d($C_1C_2T_3C_4T_5G_6$*G_7*$T_8C_9T_{10}C_{11}C_{12}$)• d($G_{13}G_{14}A_{15}G_{16}A_{17}C_{18}C_{19}A_{20}G_{21}A_{22}G_{23}G_{24}$) where the -$G_6$* G_7*- site has been modified by cisplatin, superimposed upon the averaged droplet-simulation LFMD structure (colored by element; Pt^{2+} in magenta). Pt^{2+} and its directly bounded atoms are displayed in ball and stick mode.

5.3.2 LFMD Simulation of Interstrand Cisplatin–DNA Complexes

Table 5-2. Comparison between Selected LFMD and Experimental Pt Bond Distances and Angles and RMSD of Geometrical Data for interstrand Cisplatin–DNA Complex PDB 1A2E

method	N7–Pt–N7 (deg)	Pt–N (Å)	Pt–N' (Å)	Pt–N7 (Å)	Pt–N7' (Å)	RMSD[a] (Å)
X-ray	86.8	2.010	1.997	2.000	1.999	
DommiMOE	87.9	2.033	2.037	2.000	2.001	0.023/0.265/0.494
DL_POLY/LF	90.2	1.963	1.966	1.953	1.950	0.044/0.823/2.488

[a]RMSDs for Pt–L bond lengths, heavy atoms of $\{Pt(NH_3)_2(Gua)_2\}$ and heavy atoms of whole cisplatin–DNA complex

Just as for the simulation of the 1,2–intrastrand cisplatin–DNA adduct, the droplet simulation of the interstrand cisplatin–DNA adduct started from the X-ray configuration (PDB code 1A2E). Again, a reasonable degree of the potential energy convergence was obtained after ~2500 ps (within 0.2%, Figure A3-2, Appendix 3). After ~2500 ps, the RMSD of the heavy atoms from the starting structure fluctuated around a mean of 3.23 ± 0.14 Å in the simulation. Therefore, the average structure was generated using the last 2500 ps of a 5 ns simulation. As shown in Table 5-2, the droplet LFMD average N7–Pt–N7 angle and Pt–N distances and the root-mean-square deviation (RMSD) of Pt–L bond lengths, the heavy atoms of $\{Pt(NH_3)_2(Gua)_2\}$, and the heavy atoms of the entire cisplatin–DNA complex indicate good agreement with the experiment as does the overlay of the computed and crystal structures (Figure 5-2, left). The Pt–L bond lengths have RMSDs of 0.023 Å and the heavy atom overlays for $\{Pt(NH_3)_2(Gua)_2\}$ and for the entire complex have RMSDs of 0.265 and 0.494 Å, respectively.

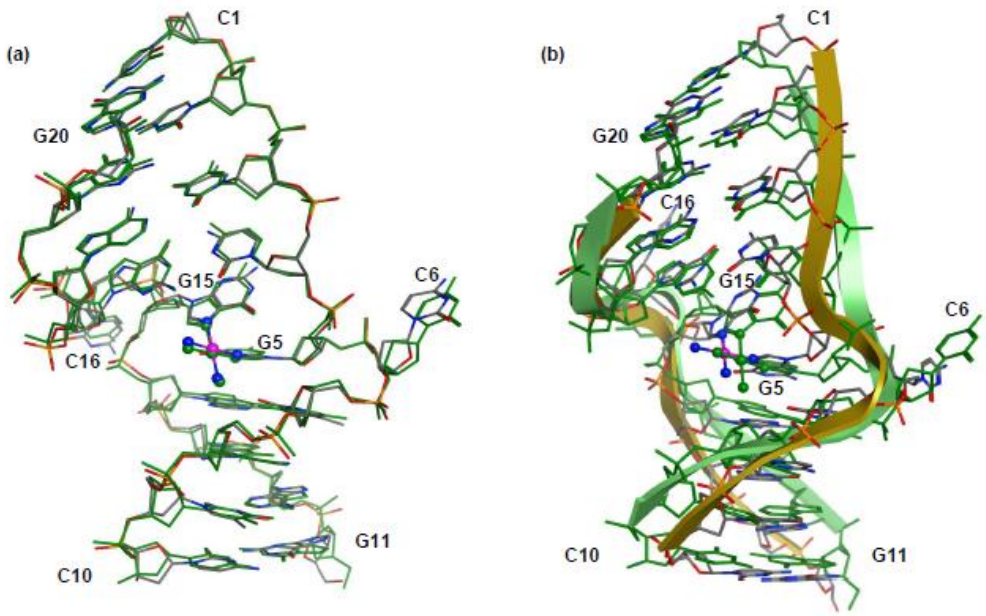


Figure 5-2. Crystal structure (green) of the cisplatin–DNA complex PDB 1A2E with the sequence d($C_1C_2T_3C_4G_5$*$C_6T_7C_8T_9C_{10}$)• d($G_{11}A_{12}G_{13}A_{14}G_{15}$*$C_{16}G_{17}A_{18}G1_9G_{20}$) where G* denotes the location of platinated nucleotides, superimposed upon the average LFMD structure (colored by element; Pt^{2+} in magenta) derived from simulations with (a) a solute molecule inside a 'drop' of solvent and with (b) periodic boundary conditions that mimic macroscopic bulk system. Pt^{2+} and its directly bounded atoms are displayed in ball and stick mode.

The RMSD for the whole system is very small leading to a concern that the droplet protocol might be introducing a surface effect where the system would have some difficulty 'escaping' from the starting structure. To check if such boundary effects lead to any artificial behaviour, a simulation of the same interstrand cisplatin–DNA complex was performed with periodic boundary conditions using our LFMM implementation in DL_POLY.

The RMSDs for Pt–L bond lengths, the $\{Pt(NH_3)_2(dGua)_2\}^{2+}$ fragment and overall heavy atoms certainly increase to 0.044, 0.823, and 2.488 Å, respectively, while the overlay of computed and X-ray structures (Figure 5-2, right) highlights the

movement of the backbone during the simulation. Compared to the periodic boundary DL_POLY simulation, the solvent droplet LFMD simulation seems to be rather 'stiffer'. Accordingly, the *cis*-$\{Pt(NH_3)_2\}^{2+}$ fragment was removed and a 6 ns MD simulation was performed to see whether the DNA would straighten out. The global bending angle does decrease from 47.2° to 24°, but the distorted duplex is unable to return to a B-DNA conformation on this time scale since one of the bases has rotated out of alignment with its complementary base (Figure A3-3, Appendix 3). This leads to the conclusion that while the droplet simulations do suffer from surface effects which tend to reinforce the starting structure, they behave correctly at a qualitative level, although droplet simulations take longer to relax than periodic boundary simulations.

5.3.3 Periodic Boundary LFMD Simulation of Interstrand P–DNA Complexes

The simulation of the interstrand **P**–DNA complex (DL_POLY2 with periodic boundary conditions) started from the modified X-ray structure (PDB entry 3CO3) as described above. The average structure generated from the last 2 ns of a 6 ns simulations is displayed in Figure 5-3a and an expanded view of the platinum-binding site is shown in Figure 5-3b. A detailed analysis of the local and global structural parameters for the nucleic acids can be generated using the program Curves+[138] as described in Section 2.6.4. The results for the **P**-DNA system are presented in Figure 5-4. For consistency and to avoid having to make comparisons involving both X-ray and NMR data which have known differences (described in Section 1.5.4.4),[21, 82] the comparison focuses on DL_POLY LFMD structures.

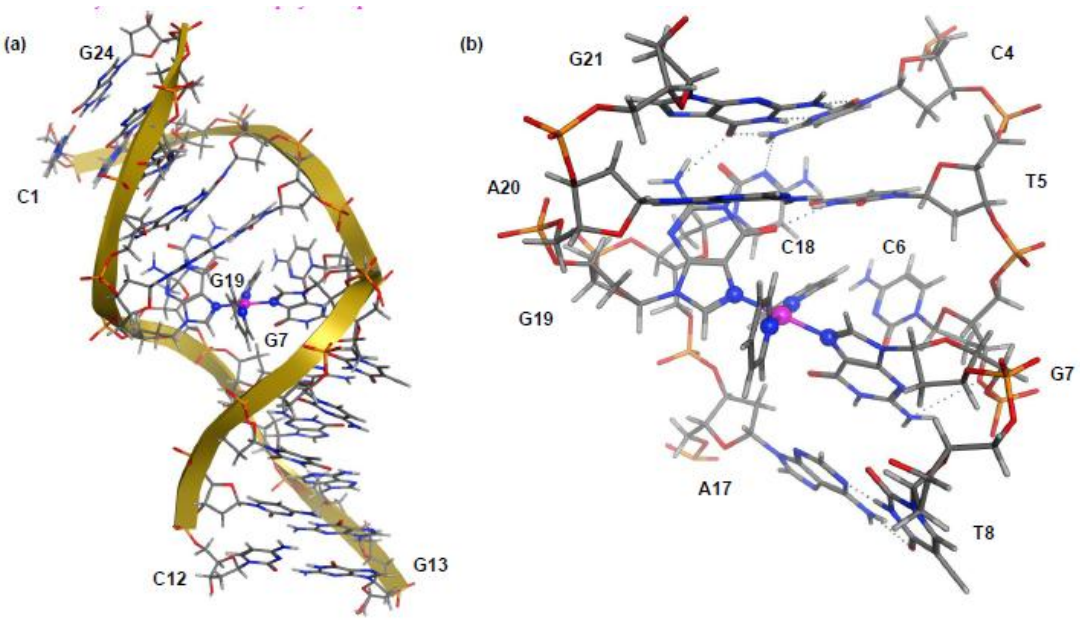
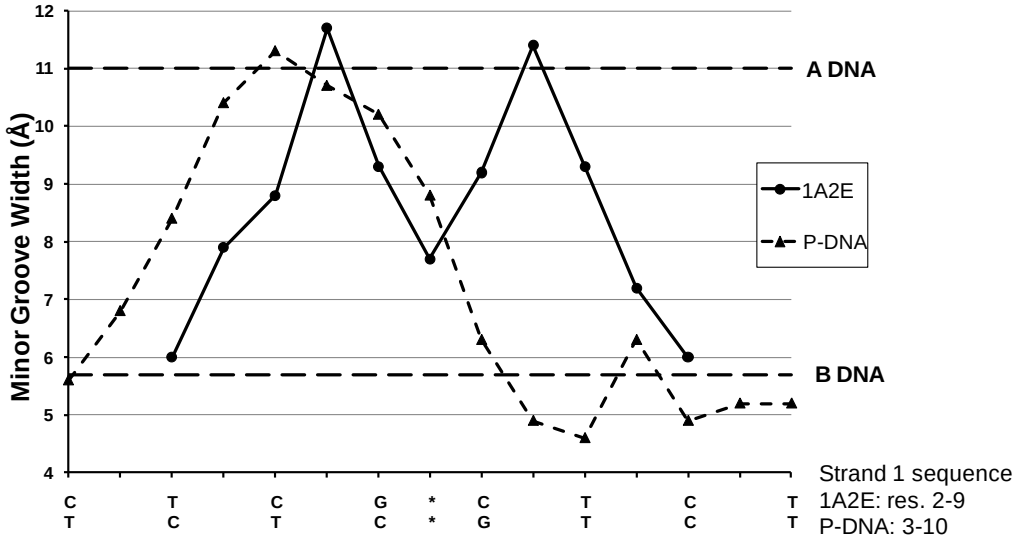


Figure 5-3. The average LFMD structure derived from simulations of interstrand **P**–DNA complex, an interstrand crosslink on the DNA duplex with the sequence d($C_1C_2T_3C_4T_5C_6G_7$*$T_8C_9T_{10}C_{11}C_{12}$)•d($G_{13}G_{14}A_{15}G_{16}A_{17}C_{18}G_{19}$*$A_{20}G_{21}A_{22}G_{23}G_{24}$) containing the *trans*-{$Pt(py)_2$}$^{2+}$ fragment coordinated to G_7 and G_{19}, with periodic boundary conditions. (a) The **P**–DNA complex. (b) An expanded view of the platinum-binding site. Pt^{2+} and its directly bounded atoms are displayed in ball and stick mode.

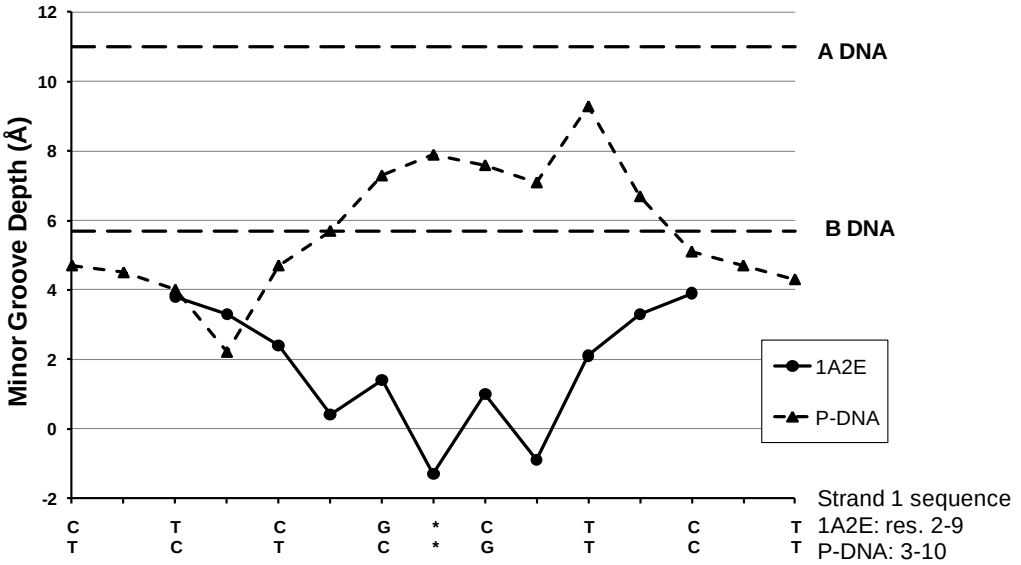
Coordination of the *trans*-platinum fragment with bulky hydrophobic pyridine ligands to G7 and G19 causes disruption of the T5-A20, C6-G19*, and G7*-C18 base pairs and a propeller twist of the platinated central base pairs by ~120°, resulting in the formation of hydrogen bonds between G19 NH2 and G21 O6, G19 O6 and A20 NH6, G21 O6 and T5 NH3, and G7 NH2 and 3' oxygen of the sugar moiety 6. On the other hand, C6 and C18 are more solvent exposed than in B-DNA, a feature similar to that found in the X-ray structure of the interstrand cross-link, PDB 1A2E (Figure 5-3b). The local distortions lead to a global bend of the duplex of ~67° and a helix unwinding of ~20°. The larger RMSD of the {$Pt(py)_2(dGua)_2$} moiety (fluctuating around a mean of 12.54± 0.33 Å) from those

in the starting structure, as compared to the RMSD of the heavy atoms for the rest of the complex (fluctuating around a mean of 10.34± 0.35), indicates a metal-binding site which is less rigid than the rest of the DNA.

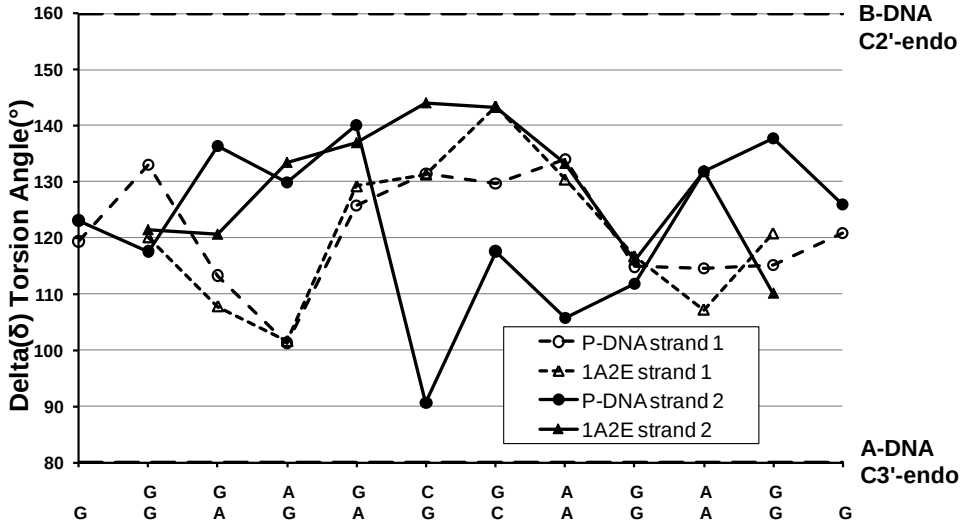
Figure 5-4a, 5-4b, 5-4c, and 5-4d summarize the minor groove width (Min-W) and depth (Min-D), C5' -C4' -C3' -O3' (δ) torsion angles, and the local inter-base pair parameters: rise, tilt, and twist for the two interstrand platinated duplexes in this work. Unlike the widened and flattered features on either side of the platination site in cisplatin–DNA complexes, the minor groove in the **P**–DNA complex is wide and shallow at one side and narrow and deep on the other. Sugar puckers can be inferred from the δ torsion angles. Such an analysis as that of Figure 5-4c shows that the **P**–DNA complex displays a more A-DNA-like C3'-endo conformation at the platinated site compared to the structure of 1A2E. The asymmetric helix unwinding on two strands of **P**–DNA complex is related to the decreasing inter-base twist angles at the metal-binding site (Figure 5-4d). A simulation with a different starting configuration may be needed in order to demonstrate the reproducibility of this structure. However, as the bulky hydrophobic pyridine ligands tend to have very few stable conformations in the core of the double helix, the possible starting configurations of **P**–DNA complex have been limited.


Minor Groove Width (Å)
12
11
10
9
8
7
6
5
4
A DNA
B DNA
1A2E
P-DNA
C T T C C T G C * * C G T T C C T T
Strand 1 sequence
1A2E: res. 2-9
P-DNA: 3-10

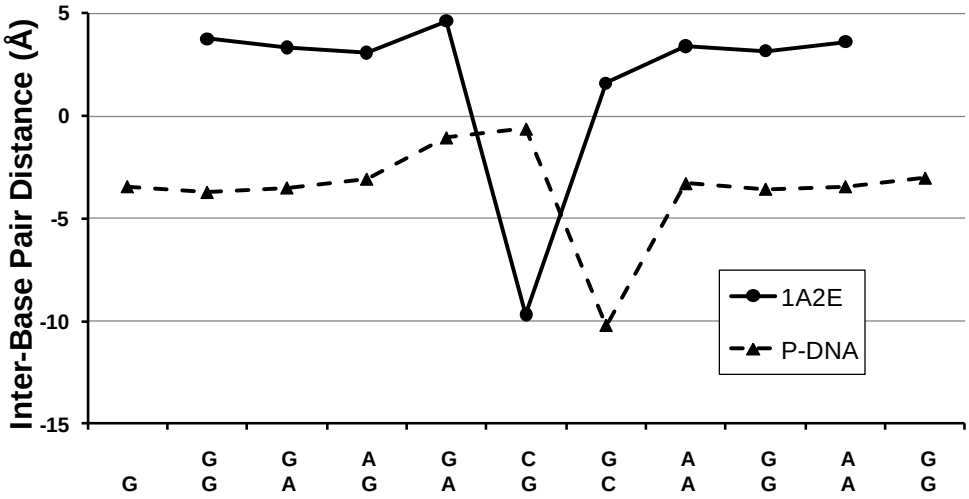
(**a**)

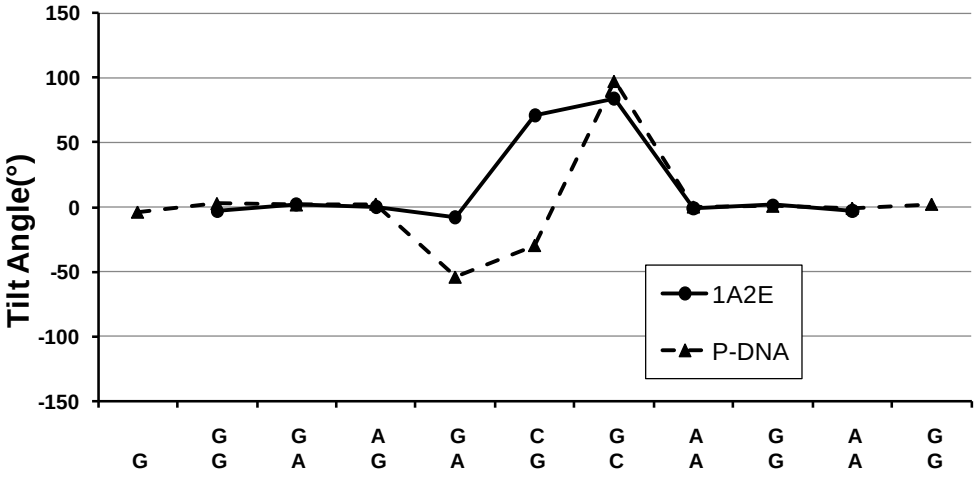

Minor Groove Depth (Å)
12
10
8
6
4
2
0
-2
A DNA
B DNA
1A2E
P-DNA
C T T C C T G C * * C G T T C C T T
Strand 1 sequence
1A2E: res. 2-9
P-DNA: 3-10

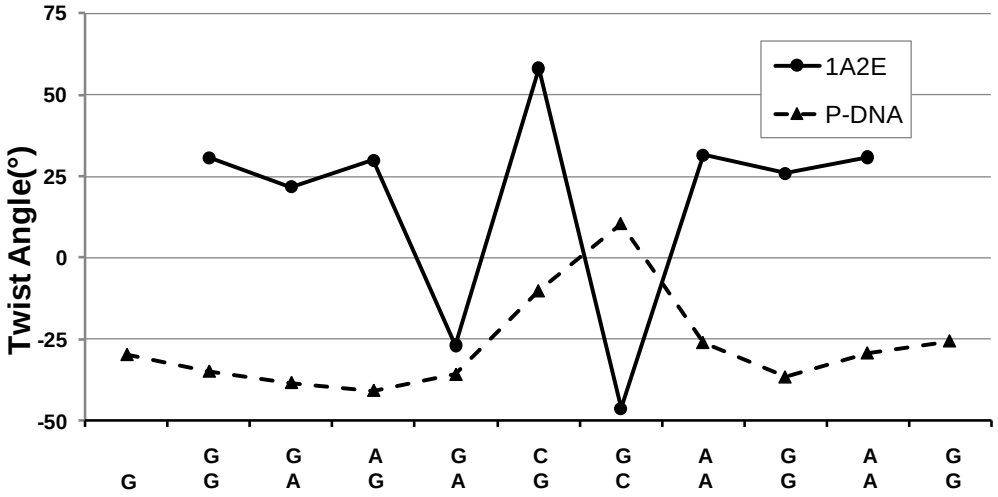
(**b**)


B-DNA
C2'-endo
A-DNA
C3'-endo
Delta(δ) Torsion Angle(°)
160
150
140
130
120
110
100
90
80
P-DNA strand 1
1A2E strand 1
P-DNA strand 2
1A2E strand 2
G G A A G A C G A A G G

(c)


Inter-Base Pair Distance (Å)
5
0
-5
-10
-15
1A2E
P-DNA
G G A A G A C G A A G G


Tilt Angle(°)
150
100
50
0
-50
-100
-150
1A2E
P-DNA
G G A A G A C G A A G G



(**d**)

Figure 5-4. Graphical representation of base pair parameters from Curves+ analysis as a function of base pair for the two of interstrand platinated duplexes. Sequences are presented from residue 2 to 9 and residue 3 to 10 in strand 1 for 1A2E and P-DNA complex, respectively, to provide an illustrative comparison between structures with different DNA sequences in (a) and (b); whereas the complementary sequences are presented from residue 20 to 11 and residue 24 to 13 in strand 2 for 1A2E and P-DNA complex, respectively in (c) and (d).

5.4 Discussion

5.4.1 Structures of the Platinated Duplex

The most interesting features of the structural distortions induced in DNA by platination are the global bend of the double helix and the deformation of the minor groove. Inspection of the overall helix bend angle and the variation in minor groove width using the program Curves+[138] reveals close agreement between LFMD and crystal/NMR structures, although this is in part due to the droplet simulation protocol. The LFMD simulations yield an overall helix bend angle of 82.0° for the 1,2–intrastrand cisplatin–DNA adduct and 67.4° (47.2° with periodic boundary conditions) for the interstrand DNA cross-link, similar to those

derived from experiments (83.2° and 60.0°, respectively).[9,10] The calculated structure of the intrastrand cross-link adduct shows that the minor groove is opened to 7.1–10.7 Å over five base pairs, from C4–G21 through T8–A17, in agreement with the respective NMR structure (7.1–11.2 Å), which is closer to the 11.0 Å width of canonical A–DNA than the 5.7 Å width of B–DNA. The average LFMD structure of the interstrand cross-link adduct gives minor groove widths of 6.5–11.2 Å (6.0–11.7 Å with periodic boundary conditions and 6.1–11.1 Å experimentally) at both ends of the platinated site, larger than the respective values of B–DNA.

Cisplatin–DNA complexes are stabilized by the formation of a hydrogen bond between the ammine-hydrogen of cisplatin and either the O=C6 group of guanine or the O=C2 moiety of thymine or an oxygen atom on the 5' -phosphate group. In contrast, the pyridine ligands of the **P**–DNA complex have aromatic C–H donors and aromatic π-acceptors which may interact with aromatic groups on proteins. for C–H$\cdots\pi$-interactions. These weak C–H$\cdots\pi$ hydrogen-bond-like interactions each provide a stabilization energy of about 0.5–1.0 kcal/mol and may thus play an important role in protein–DNA recognition.[89] Another notable feature of the **P**–DNA complex is that the weak stabilization due to π–π stacking between one of the hydrophobic pyridines and T5 results in a relatively flexible local metal-binding site, which could increase the specificity of protein binding for the platination site.

Inevitably, how well platinum–DNA complexes can be modeled is based on the interplay between LFMM for the metal centre and conventional MM for the 'organic' parts. For the latter, the performance of the most popular nucleic acid force fields has been evaluated (refer to Section 1.5.3 for the summary) and, in

general, simulation is capable of reproducing a variety of nucleic acid structures. Here, I acknowledge that the performance with respect to describing intricate structural parameters of platinum–DNA complexes using LFMM in conjunction with various nucleic acid force fields should be constantly reviewed.

Brabec *et al.* subsequently demonstrated experimentally that, in addition to 5'-CG/5'-CG crosslinks like the **P**–DNA complex modeled in the present study, 5'-GC/5'-GC crosslinks and crosslinks between G and the complementary C are formed when complex **17** is photoactivated in the presence of DNA. It will be interesting to model the latter crosslinks in future work.

5.4.2 Biological Significance

The main thrust of this work has been to probe the effect on the structure of DNA of generating a *trans*-$\{Pt(py)_2\}^{2+}$ interstrand crosslink arising from the photoreduction of the potent cancer-cell-cytotoxic complex *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ (**17**). The structure of the *trans* interstrand **P**–DNA complex generated from our simulations exhibits more bending (67°) and less unwinding (20°) than an interstrand cisplatin–DNA adduct (bent by 60° and unwound by 81°) as well as less bending than a reported transplatin interstrand cross-link (bent by 26° and unwound by 12°).[177] The proposed photoadduct bears a closer resemblance to the intrastrand cisplatin–DNA adduct, PDB 1A84, (bent by 83.2° and unwound by 36°) and is of potential importance in the molecular mechanism of action of this photoactivatable complex.

A quantitative analysis of the influence of nucleosomal core proteins on specific drug–DNA interactions using a *Taq* DNA polymerase stop assay has revealed that larger platinum compounds generally display a greater tendency to target the

linker region of the nucleosomal DNA and have less access to nucleosomal core DNA.[182] For example, dichloro(N-[3-[(2-aminoethyl)amino]propyl]-9-aminoacridine-4-carboxamide)platinum(II) with a molecular weight of 603 produces a higher average linker-core ratio (L/C ratio) of 2.52 compared to cisplatin (molecular weight of 300, L/C ratio of 1.32). Upon binding to the linker region of the nucleosomal DNA, the platinum modification could (i) enhance the binding of histone H1 and the high mobility group (HMG1) protein[185] and shield the damage site from recognition by DNA repair machinery[6,177,185] by having a dramatic effect on the global curvature of the DNA similar to a cisplatin 1,2-intrastrand lesion, increasing the flexibility of the helix at the local site of the interstrand cross-link, (ii) attenuate the negative electrostatic potential of backbone phosphates by widening the minor groove at one end of the platinated site, and (iii) change the groove landscape by deepening the minor groove at the opposite end of the platinated site together with A-DNA-like conformations and the above structural factors, and thus affect the formation of base-specific hydrogen bonds in the major groove and the electrostatic potential in the minor groove, two general mechanisms for protein–DNA recognition.[186] The compound can also access nucleosomal core DNA. With a striking similarity to the global shape of the intrastrand *cis*-$\{Pt(NH_3)_2\}^{2+}$ 1,2-d(GpG) cross-link, the *trans* interstrand **P**–DNA adduct may influence the rotational and translational positioning of DNA in nucleosomes.[180]

In summary, photoactivation promotes the rapid formation of a *trans* interstrand cross-link. Various repair proteins that bind preferentially to platinum intrastrand cross-links do not interact with interstrand cross-links. DNA interstrand cross-links, in general, prevent strand separation, and thus act as a physical barrier to

DNA replication and/or DNA transcription, which, in contrast to G2 arrest in the cell cycle caused by cisplatin damage,[6] may block advance through the S phase and drive the cells effectively into apoptosis. In addition, the distinctive local and global shape and the dynamic features of the **P**–DNA complex may promote a higher affinity for histone H1 and HMG1, which can shield the platinum adducts from DNA repair enzymes such as nucleotide excision repair (NER) proteins, which mainly recognize and excise intrastrand adducts, and increase the cytotoxicity of anticancer agents. Preclinical studies have confirmed that modulation of repair pathways can enhance the sensitivity to DNA damaging agents.[83] Moreover, the protein-binding profile of the proposed *trans* interstrand **P**–DNA complex may be very different from clinically effective platinum drugs, and may modulate the binding and assembly of transcription factors such as the TATA-binding protein (TBP). TBP recognizes the TATA box of the promoter by intercalation between four phenylalanine residues and base pairs of the TATA box in the minor groove causing substantial unwinding and bending of DNA. It binds to cisplatin adducts with similar affinity and kinetics compared to TATA boxes,[187] and therefore affects gene expression. The targeting to cancer cells by site-specific photoactivation, reduced DNA repair, cell-cycle arrest in the S phase, and the influence on gene expression may all contribute to the high potency and unusual mechanism of action of complex **17** with potential for circumventing resistance to current Pt drugs.

5.5 Conclusions

The ligand field molecular mechanics (LFMM) parameters for Pt(II) complexes have been extended to the interaction between the N7 position of guanine to

enable ligand field molecular mechanics dynamics (LFMD) simulations of the intrastrand and interstrand cisplatin–DNA adducts as well as a **P**–DNA complex, a *trans*-$[Pt(py)_2(Gua)_2]^{2+}$ interstrand cross-link formed by binding of the major photoproduct of a novel platinum(IV) complex *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ (**17**) to two nucleobases in DNA. The main findings of this work are as follows.

(i) LFMM yields metal-ligand bond distances in good agreement with the experimental and QM/MM-optimized small-molecule structures, while LFMD accurately reproduces the X-ray and NMR structures of the much larger intrastrand and interstrand platinum–DNA complexes at a significantly reduced computational cost compared to QM/MM.

(ii) The unusual nature of the proposed *trans* interstrand cross-link in **P**–DNA complex promoted by photoactivation of *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ (**17**), compared to the predominantly 1,2-intrastrand cross-links formed by cisplatin, may contribute to the former's high potency towards cancer cells.

(iii) Unlike the widened and flattened features in cisplatin–DNA complexes, the minor groove in the **P**–DNA complex is widened at one end and deepened at the other end of the platinated site. The distinctive structure of the *trans* interstrand **P**–DNA complex with more bending (67°) and less unwinding (20°) is closer to intrastrand cisplatin–DNA adducts (bent by 83.2° and unwound by 36°) than interstrand cisplatin–DNA adducts (bent by 60° and unwound by 81°) or transplatin interstrand cross-links (bent by 26° and unwound by 12°). The analysis of the δ torsion angles indicates a more A-DNA-like C3'-endo conformation at the platinated site for the **P**–DNA complex than for the interstrand cisplatin–DNA complex (PDB 1A2E). The combination of the structural features described here

offers numerous possibilities for specific interactions with proteins that may help rationalize to the striking differences in cellular repair mechanisms and apoptosis pathways and influence the loading of DNA and RNA polymerases in the promoter region of genes.

Chapter 6

Perspectives and Future Work

Herein, the major advances and conclusions in computer-aided drug design of photoactivated platinum complexes within this thesis are summarised. Future work for a better understanding of the mechanism of action as well as efficient drug design and discovery to achieve greater potency and desirable pharmacological properties is outlined.

6.1 Concluding Remarks

A variety of platinum complexes have proven successful as anticancer agents in the clinic and in clinical trials. In spite of revolutionary leaps forward in understanding cancer pathogenesis, relatively little progress has been made in exploiting these insights into etiology (i.e. the causative mechanisms of disease) to prevent the disease and, equally important, to treat it. Most of the anticancer treatments in widespread use today were developed prior to 1975, at a time when the development of therapeutics was not yet informed by the genetic and biochemical mechanisms of cancer pathogenesis. Death rates for cancer have begun to fall, because of early detection and surgical removal of growths that have advanced through only the early stages of tumour progression. However, mortality caused by the more advanced tumours that have invaded and begun to metastasize has changed little in recent decades.[188] Among a number of strategies of therapy, photoactivated platinum compounds offer both temporal and spatial control over drug activation and have remarkable potential to increase selectivity for cancer cell lines and to reduce some of the noxious side-effects associated with conventional chemotherapeutics.[35,36,159]

Numerous experimental and theoretical studies have been carried out on metal-based drugs and their *in vivo* mechanisms of action.[6,11,31,32,40,189-191] However, the

research on photoactivated metal complexes is still in its infancy.[35,36,159,161,192,193] Many important issues such as the principles that govern the power and frequency of radiation, structures of metal complexes, properties in excited states, and the induced photopathways and photoproducts as well as solvent effects are not fully understood.

In this thesis, DFT/TDDFT has been used as a tool to help tune the coordination environment of platinum(IV) complexes to achieve desirable photochemical properties and furthermore LFMM/LFMD has been employed to study novel distortions induced by DNA platination which are important for understanding the mechanism of action of a new class of photoactive platinum anticancer drugs. In Chapter 3, exploration of *trans* influences, the nature of OR^- and (pseudo)halogen ligands, electron-withdrawing/donating/delocalizing substituents on the N-ligands, and intramolecular H-bonds by TDDFT shows that the design of platinum(IV) complexes with absorption bands shifted towards longer wavelengths (~330 nm) can be achieved by introducing intramolecular H-bonds involving the OH ligands and 2-hydroxyquinoline or by iodido ligands, and the direction of metal-ligand electron transfer can be controlled. Furthermore, mesomeric electron-withdrawing substituents on pyridine result in low-energy absorption with significant intensity in the visible region. The comparison of the optimised structures shows that, in general, Pt(IV) complexes with Pt−L bond distances closer to their Pt(II) products are more likely to undergo photoreduction with longer-wavelength light. The molecular orbitals analysis shows that the distinct makeup of the molecular orbitals in electronic transitions for *cis*/*trans*-{$Pt(N_3)_2$} isomers result in different photoproducts.

Ligand field molecular mechanics (LFMM) parameters for Pt–guanine interactions are derived to enable the square-planar geometry of the platinum fragment to be generated automatically rather than be imposed by artificial forces (see Chapter 4), and are applied to the first ligand field molecular dynamics (LFMD) simulations of bifunctional DNA adducts of the anticancer drug cisplatin to validate our protocols before advancing to a possible major DNA photoproduct of the novel platinum(IV) complex *trans,trans,trans*-$[Pt(N_3)_2(OH)_2(py)_2]$ (**17**) (py = pyridine) (see Chapter 5). LFMM gives an excellent description of the local Pt coordination which is in good agreement with experimental and QM/MM structures. LFMD also accurately reproduces the experimentally-observed structures of intrastrand and interstrand platinum–DNA complexes. The major photoproduct from complex **17** can form an interstrand crosslink on the DNA duplex dodecamer, $d(C_1C_2T_3C_4T_5C_6G_7{*}T_8C_9T_{10}C_{11}C_{12})\bullet d(G_{13}G_{14}A_{15}G_{16}A_{17}C_{18}G_{19}{*}A_{20}G_{21}A_{22}G_{23}G_{24})$, where the *trans*-$\{Pt(py)_2\}^{2+}$ fragment is coordinated to G_7 and G_{19}. This adduct has a widened minor groove at one end and deepened minor groove at the opposite end of the platinated site, and exhibits a global bend of ~67° and an unwinding of ~20°. The unusual structural features associated with this interstrand *trans*-$[Pt(py)_2(Gua)_2]^{2+}$ cross-link offer many possibilities for specific protein–DNA interactions, which suggests possible mechanisms for the high potency of this photoactivated complex. LFMM/LFMD appears to provide a reliable way to model local and global conformational distortions induced by binding of various platinum complexes, particularly highly strained systems involving critical conformational changes. It will be interesting to model the other forms of interstrand crosslinks using LFMD in future work.

Both experimental and theoretical works suggest that the photochemistry of platinum complexes is complex and thus not simple to predict, and that photoproducts can be complicated to characterise, and must therefore be investigated thoroughly to elucidate the pathways. Future work will focus on applying LFMD to the study of sequence-dependent reactions of DNA with platinum-based anticancer drugs and their photoproducts, and to the design of metal-containing complexes that exhibit specificity with respect to duplex DNA and also between different G-quadruplexes.

6.2 Future Work: Rational Design of Metal-Based Complexes with Selectivity for Duplex and Quadruple DNA

In order for drug targeting to be selective and effective, it is not enough for metal-based complexes to bind to duplex or quadruplex DNA. New metallopharmaceuticals must also target selectively one duplex/quadruplex over another. LFMD simulations could be used for the docking of metal complexes to specific DNA sequences and structures and to compute relative kinetic and thermodynamic free energies for the binding of metal complexes to DNA with desirable sequences and structures, as well as to predict transition structures of metal–DNA adducts and DNA conformational changes upon metal binding.

In DNA-drug docking, the drug is brought (manually or, in some cases, automatically) into the vicinity of the binding site and oriented so that electrostatic and van der Waals interactions (computed via appropriate molecular mechanics energy expressions) between the drug and DNA are optimized. Accurate docking calculations can be very difficult and require an extremely fast computer because both drugs and nucleic acids are relatively flexible molecules that mutually adjust

their conformations in the process of binding to maximize steric and chemical complementarity. Alternatively, to speed up the calculation, one or both components can be frozen. For example, the active site can be fixed and a set of precalculated drug conformations docked to give a series of 'poses' each of which has a 'score'. Such calculations are very quick but potentially of low accuracy.

Ideally, I would like to compute the binding free energy. Free energy calculations using molecular dynamics or Monte Carlo simulations provide a direct link between the microscopic structure and fluctuations of the model system and its most important macroscopic thermodynamic property, and thus show great utility for guiding molecular design. Free-energy perturbation and thermodynamic integration calculations provide rigorous means to compute free-energy changes so that accurate prediction of metal-biomolecule binding affinities can be achieved.[194] The fundamental difficulty in free energy calculations lies in obtaining adequate sampling of conformations. Because of the ruggedness of the energy landscapes of nucleic acids and proteins, many energy barriers cannot be crossed spontaneously in simulations spanning even a few nanoseconds. In addition, due to the inherently very short elementary simulation time step, ~2 fs, molecular dynamics simulations are presently limited to relatively short total simulation times, on the order of nanoseconds to microseconds. However, with the ongoing significant advances in hardware and software, as well as techniques to enhance sampling, the evaluation of the kinetic and thermodynamic reactivity of metal–DNA interactions could become a powerful tool in the design and synthesis of new potent metallopharmaceuticals with sequence-dependent selectivity for duplex DNA and with accurate targeting of the loops and grooves of G-quadruplexes. In addition, the characterization of transition states is essential if the

chemical processes are under kinetic control. As LFMM employs Morse potential to model metal-ligand bonds, transition states of metal-ligand interactions can be computed if proper parameters, such as the change of charges, are provided.

6.3 Future Work: the Study of the Interactions between Various Metal Ions and DNA/Proteins

Not only do platinum complexes bind to DNA, but they can also interact with proteins involved in replication, transcription, repair, and chromatin structure. Many of these proteins communicate with each other to mediate drug uptake, signal transduction, and gene expression, which further affect anticancer activity of platinum-based drugs. Better information about the structures and the reactivity of platinum complexes with a variety of proteins will be valuable for elucidating the molecular mechanism of these compounds and further improving drug design. On the other hand, LFMM force field can be extended to adenine and other transition metals, such as Ruthenium, Osmium, Rhodium, Iridium, and Gold. Careful parameterisation is required to ensure that the preference of guanine over adenine and of AG 1,2-intrastrand crosslinks over the analogous GA adducts, and periodic trends of metal-ligand bonds can be reproduced.

Appendices

Appendix 1

Table A1-1. Comparison between Calculated and Experimental Bond Distances (in Å) and Angles (in deg) for *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$ and *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$

Complex		BP86/DZT[a]	BP86/TZP[b]	Experiment
trans, trans, trans-$[Pt(N_3)_2(OH)_2(NH_3)_2]$	Pt-N_3	2.088	2.090	2.046
	Pt-NH_3	2.051	2.057	2.036
	Pt-OH	2.050	2.047	2.006
	Pt-N-N	116.3	117.2	117.2
	N-N-N	175.6	174.7	174.5
cis, trans, cis-$[Pt(N_3)_2(OH)_2(NH_3)_2]$	Pt-N_3	2.052	2.053	2.037
	Pt-NH_3	2.108	2.116	2.033
	Pt-OH	2.049	2.046	2.006
	Pt-N-N	115.5	116.3	116.3
	N-N-N	175.5	174.8	172.9
	Rmsd (Å)	0.044	0.046	
	Rmsd (°)	1.5	0.9	

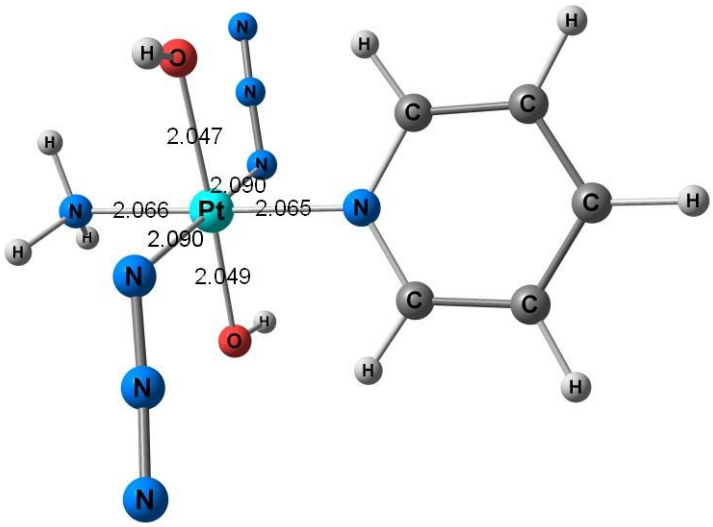
[a] BP86 functional in combination with the uncontracted triple-ζ+polarisation (TZP) STO basis sets. [b] BP86 functional in conjunction with the uncontracted double-ζ+polarisation (DZP) STO basis sets.

Table A1-2. Comparison of intense bands λ_{max} between calculated electronic transitions with scalar and two-component (spin-orbital coupling) relativistic TDDFTs using SAOP potentials and experimental UV-Vis spectrum for *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$, *cis, trans, cis*-$[Pt(N_3)_2(OH)_2(NH_3)_2]$, and *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(NH_3)(pyridine)]$ (**1**).

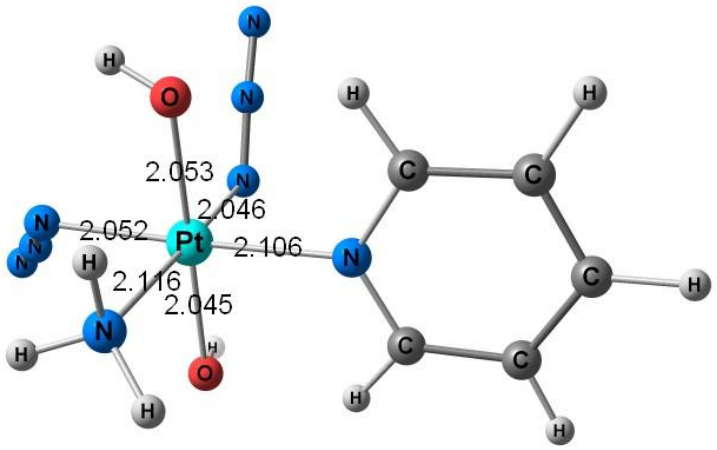
trans, trans, trans-$[Pt(N_3)_2(OH)_2(NH_3)_2]$						
	nm	1/cm	eV	kcal/mol	transition	Oscillator Strength
experiment	285	35088	4.35	100.32		
scalar	295	33924	4.21	96.99	11	0.341
	237	42245	5.24	120.79	19	0.192
scalar + COSMO[a]	301	33269	4.13	95.12	10	0.218
	282	35419	4.39	101.27	11	0.207
	225	44508	5.52	127.26	19	0.231
two-component	293	34104	4.23	97.51	47	0.313
cis, trans, cis-$[Pt(N_3)_2(OH)_2(NH_3)_2]$						
	nm	1/cm	eV	kcal/mol	transition	Oscillator Strength
experiment	256	39063	4.84	111.68		
scalar	277	36046	4.47	103.06	13	0.134
	271	36944	4.58	105.63	14	0.156
two-component	277	36083	4.47	103.17	55	0.114
trans, trans, trans-$[Pt(N_3)_2(OH)_2(NH_3)(pyridine)]$ (**1**)						
	nm	1/cm	eV	kcal/mol	transition	Oscillator Strength
experiment	289	34602	4.29	98.93		
scalar	312	32025	3.97	91.57	17	0.119
	293	34112	4.23	97.53	21	0.261
two-component	311	32200	3.99	92.07	71	0.115
	292	34280	4.25	98.01	87	0.225

[a] scalar relativistic TDDFT in combination with COSMO.

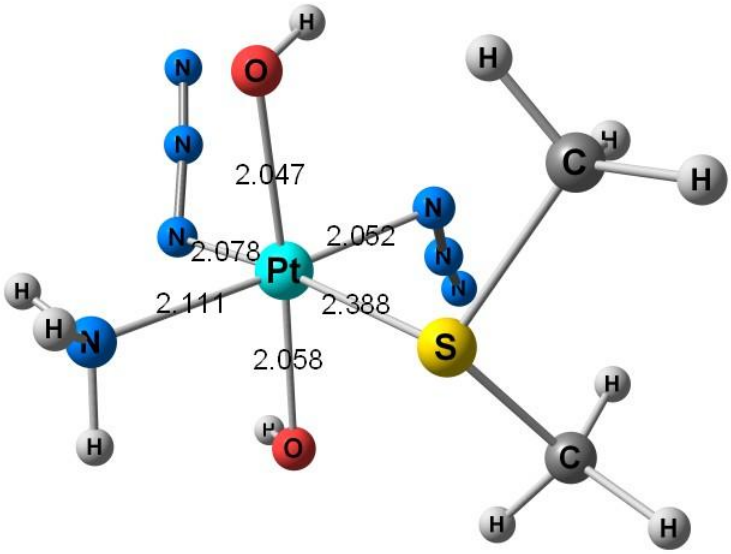
1


2.047
2.090
2.066
Pt
2.065
2.090
2.049

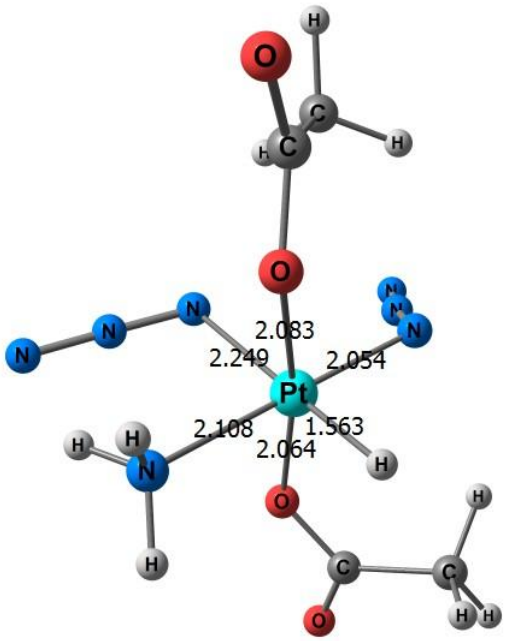
2


2.053
2.046
2.052
Pt
2.106
2.116
2.045

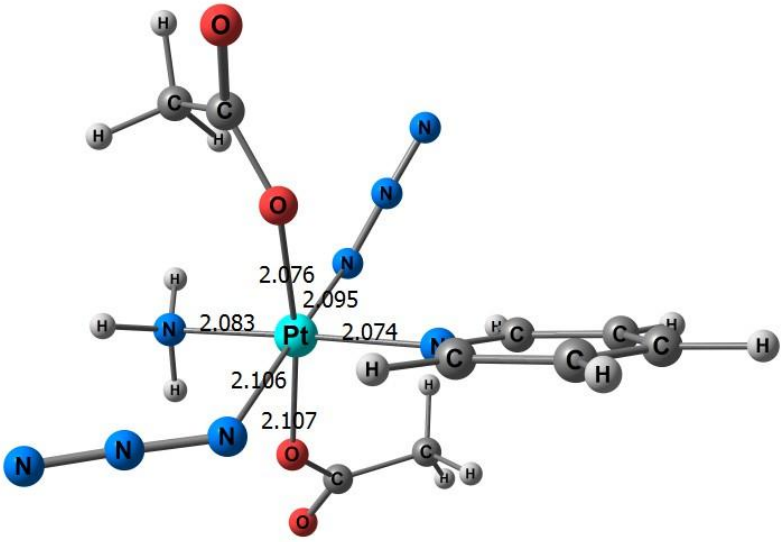
3


2.047
2.078
Pt
2.052
2.111
2.388
2.058

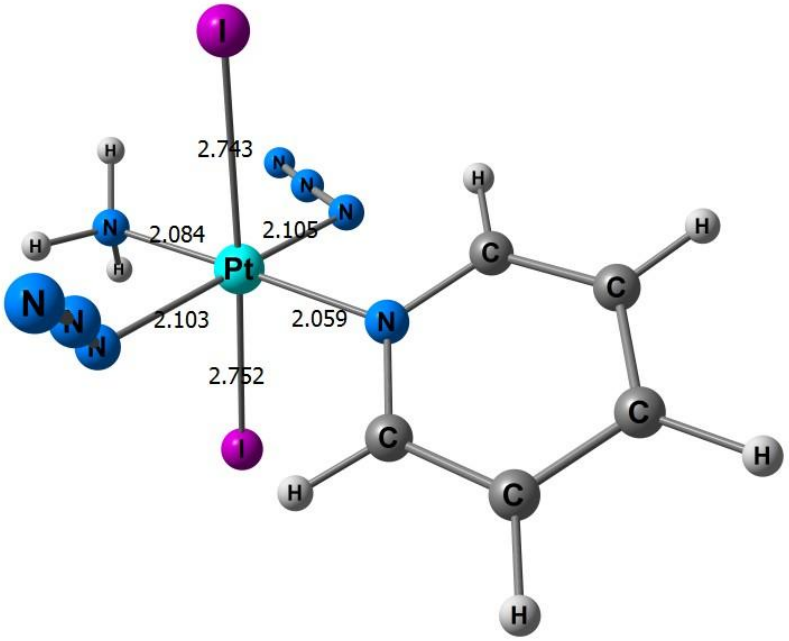
4



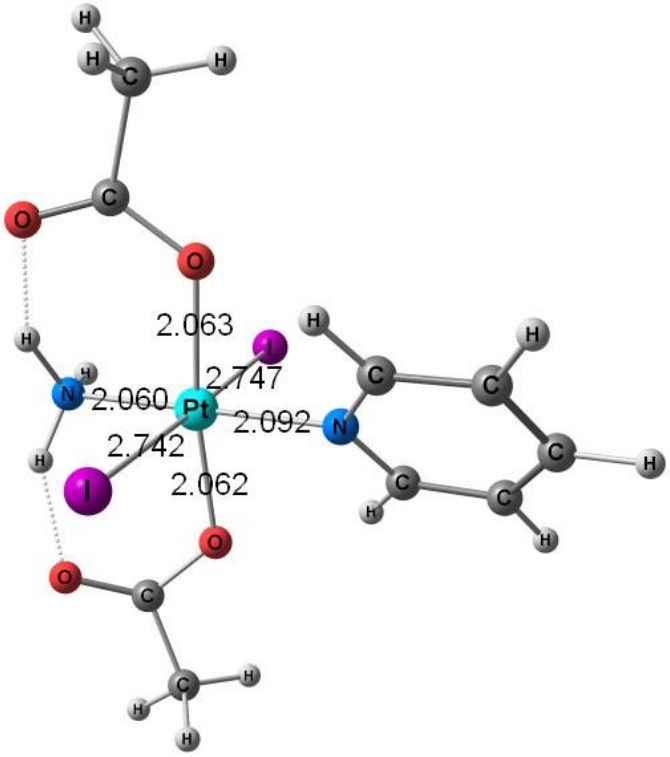
5



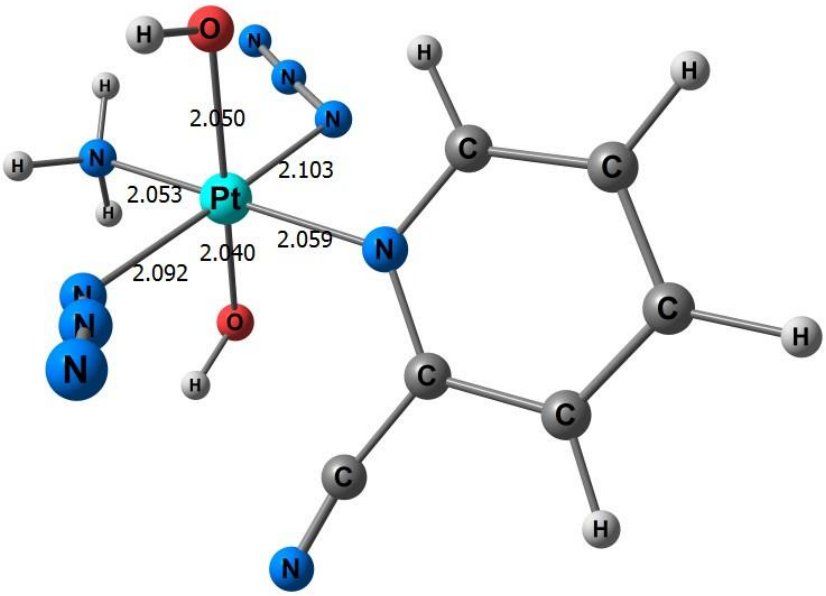
6


I
2.743
N
N
N
H
N
H
2.084
2.105
Pt
N
N
N
2.103
2.059
N
2.752
I
C
C
H
C
C
H
H
C
H
H

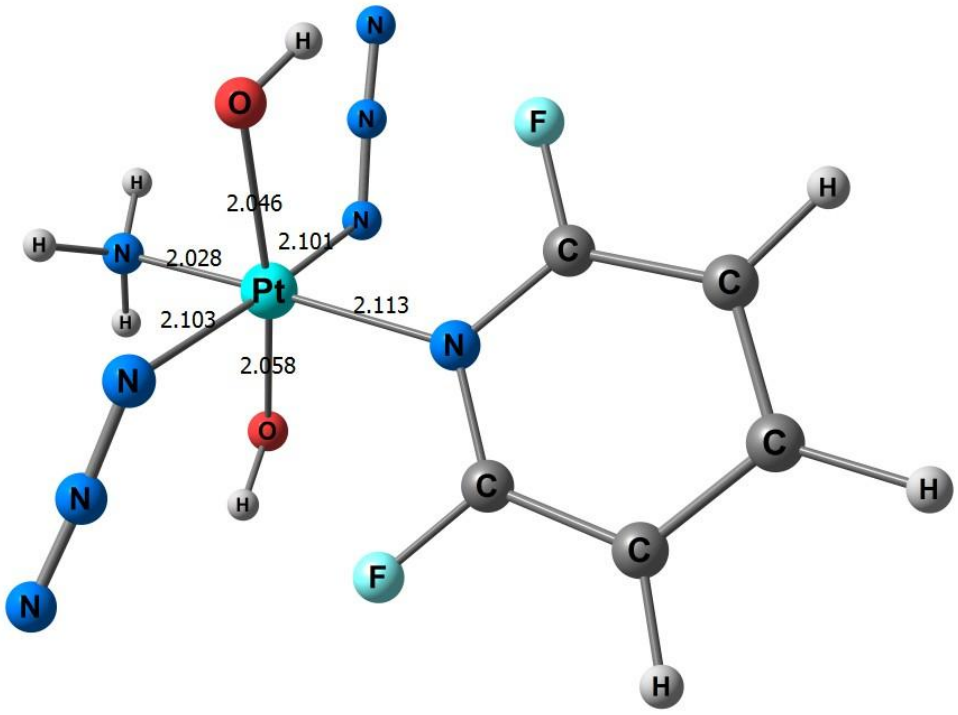
7


H
H
C
H
O
C
O
2.063
I
H
N
2.060
Pt
2.747
2.092
N
2.742
I
2.062
C
C
H
C
C
H
H
O
O
C
C
H
H
H

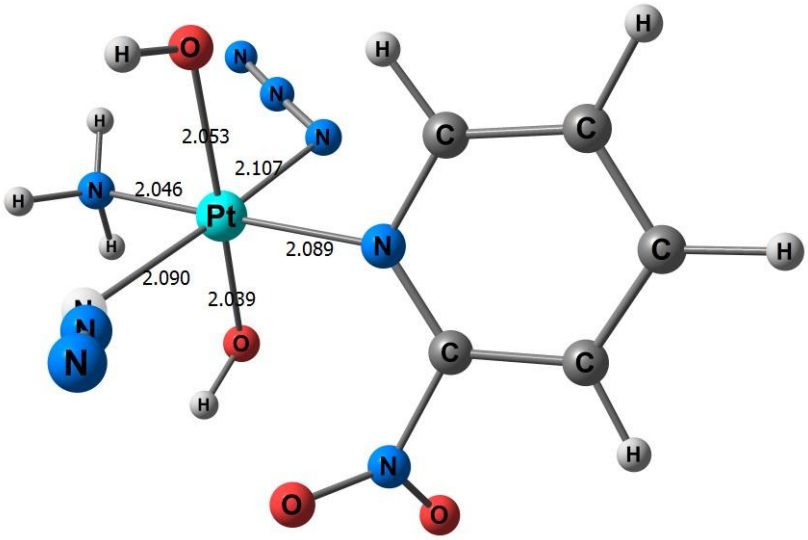
8


2.050
2.103
2.053
Pt
2.059
2.040
2.092

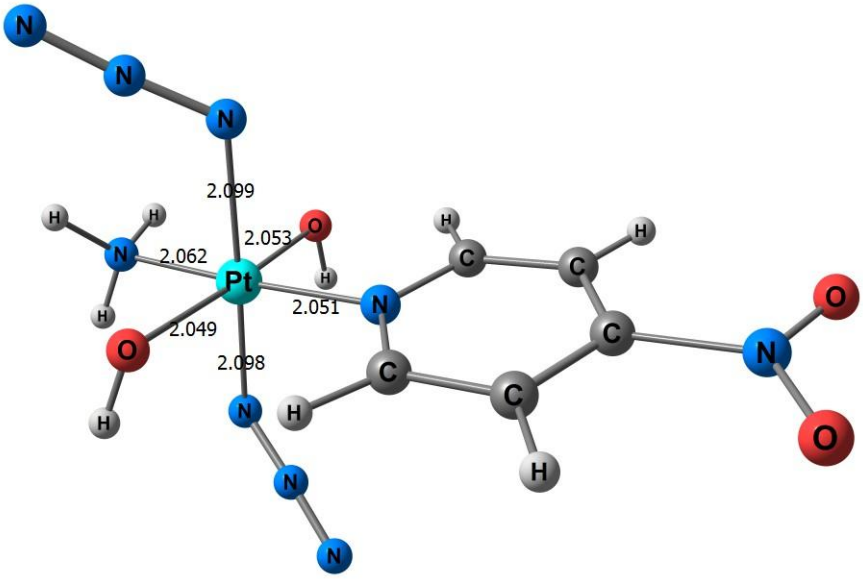
9


2.046
2.101
2.028
Pt
2.113
2.103
2.058

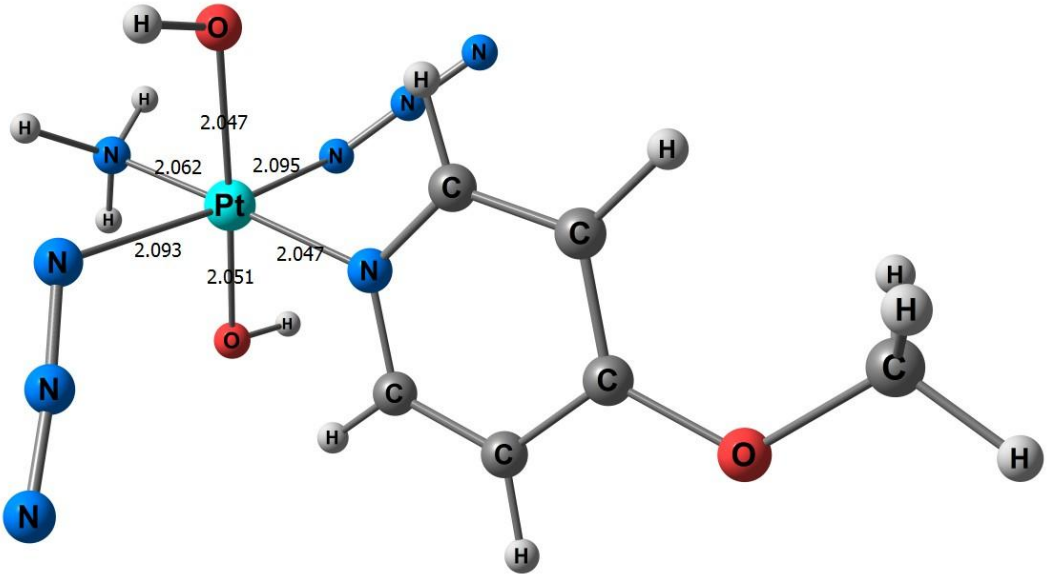
10


2.053
2.107
2.046
2.089
2.090
2.039

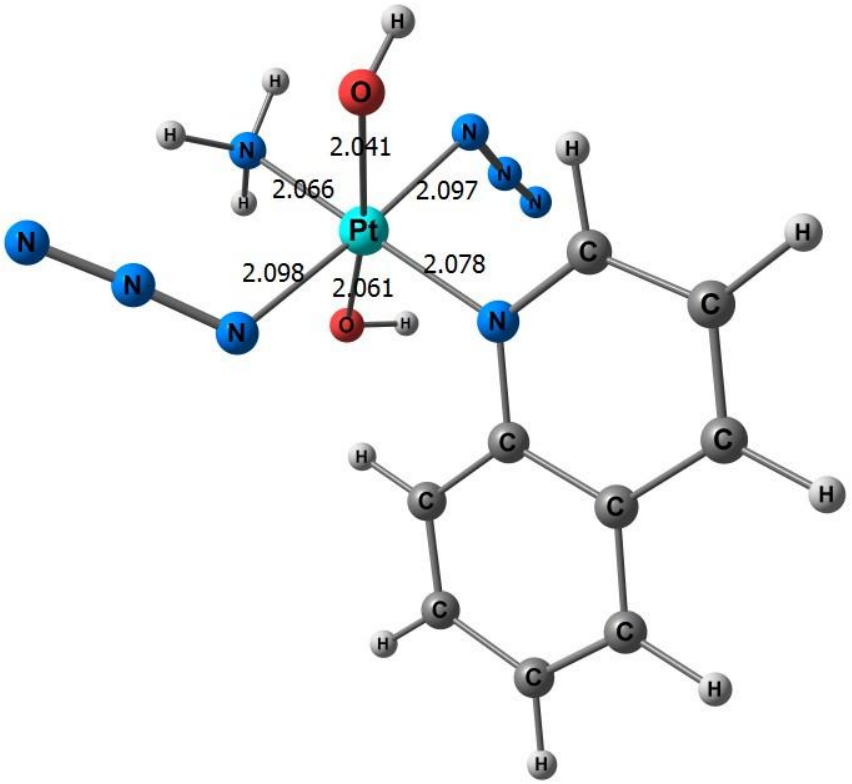
11


2.099
2.053
2.062
2.051
2.049
2.098

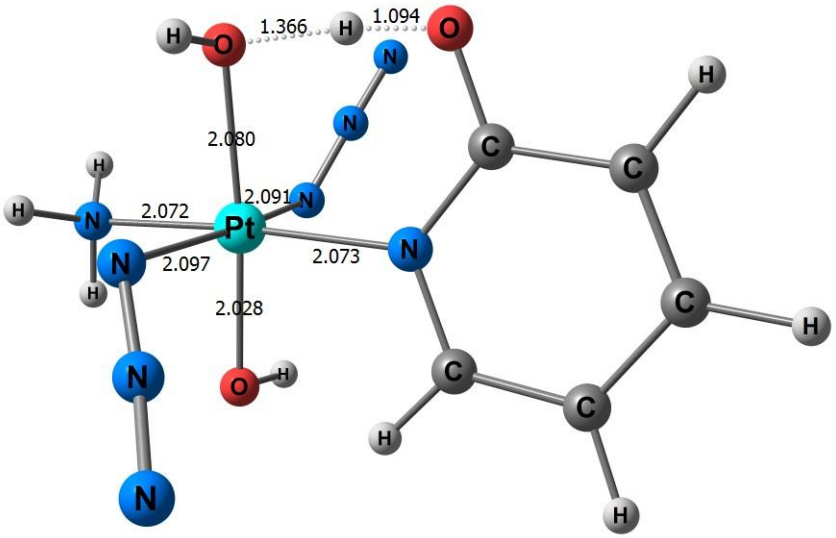
12


2.047
2.062
2.095
2.093
2.047
2.051

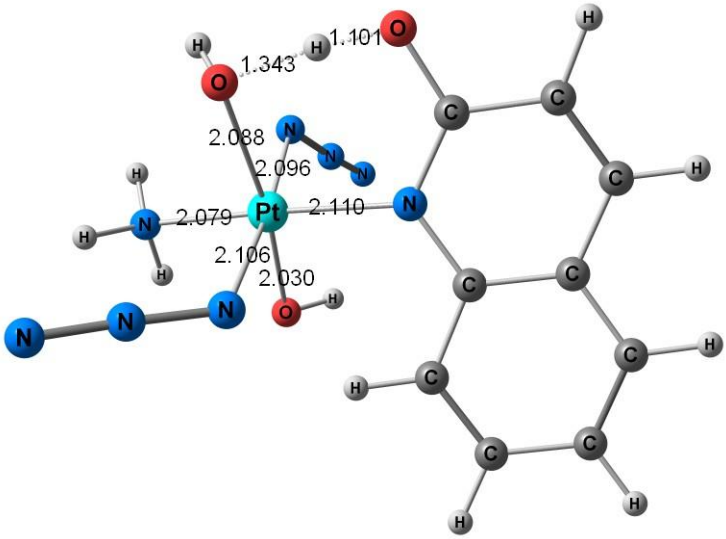
13


H
O
2.041
N
H
N
2.066
2.097
N
N
Pt
2.098
2.078
2.061
N
O
H
N
N
C
H
C
C
H
C
C
H
C
C
C
H
C
H

14


H
O
1.366
H
1.094
O
N
N
2.080
C
H
C
2.091
N
H
N
2.072
Pt
N
2.097
2.073
C
H
H
2.028
C
N
O
H
C
H
N
C
H

15



16

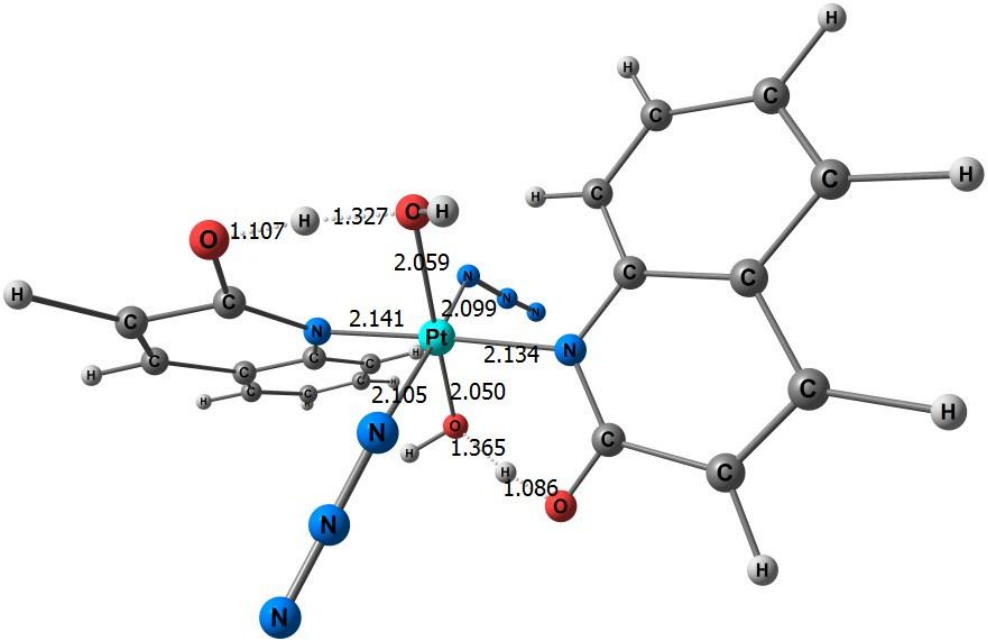


Figure A1-1. Fully optimized structures of the platinum(IV) complexes. Bond lengths are given in Å .

Appendix 2

Table A2-1. New LFMM/MMFF94 and LFMM/AMBER94 charge schemes for dGuo(Pt)$^+$ derived from CHelpG charges of *cis*-[Pt(NH$_3$)$_2$(Cl)(dGuo)]$^+$.

		QM				mmff94			new	new
Label	Atom	dGuo	dGuo(Pt)+	(dGuo(Pt)+)-dGuo		dGuo	dGuo(Pt)+	(dGuo(Pt)+)-dGuo	(dGuo(Pt)+)-dGuo	dGuo(Pt)$^+$
1	N1	-0.894	-0.888	0.006		-0.490	-0.490	0.000	0.006	-0.484
2	C2	1.006	1.094	0.088		0.560	0.560	0.000	0.088	0.648
3	N3	-0.772	-0.776	-0.004		-0.576	-0.576	0.000	-0.004	-0.580
4	C4	0.402	0.548	0.146		-0.026	-0.026	0.000	0.146	0.120
5	C5	-0.026	-0.328	-0.302		0.141	0.141	0.000	-0.302	-0.161
6	C6	0.794	0.888	0.094		0.716	0.716	0.000	0.094	0.810
7	N7	-0.56	-0.147	0.413		-0.565	-0.012	0.553	0.413	-0.152
8	C8	0.235	0.168	-0.067		0.037	0.037	0.000	-0.067	-0.030
9	N9	-0.268	-0.231	0.037		0.048	0.048	0.000	0.037	0.085
10	N10	-0.944	-1.042	-0.098		-0.850	-0.850	0.000	-0.098	-0.948
11	O11	-0.609	-0.657	-0.048		-0.570	-0.570	0.000	-0.048	-0.618
12	H12	0.445	0.457	0.012		0.370	0.370	0.000	0.012	0.382
13	H13	0.388	0.458	0.07		0.400	0.400	0.000	0.070	0.470
14	H14	0.403	0.466	0.063		0.400	0.400	0.000	0.063	0.463
15	H15	0.147	0.157	0.01		0.150	0.150	0.000	0.010	0.160

						Amber94			new	new
						dGuo	dGuo(Pt)+	(dGuo(Pt)+)-dGuo	(dGuo(Pt)+)-dGuo	$dGuo(Pt)^{+}$
						-0.505	-0.505	0.000	0.006	-0.499
						0.743	0.743	0.000	0.088	0.831
						-0.664	-0.664	0.000	-0.004	-0.668
						0.181	0.181	0.000	0.146	0.327
						0.199	0.199	0.000	-0.302	-0.103
						0.492	0.492	0.000	0.094	0.586
						-0.573	-0.573	0.000	0.413	-0.160
						0.074	0.074	0.000	-0.067	0.007
						0.058	0.058	0.000	0.037	0.095
						-0.923	-0.923	0.000	-0.098	-1.021
						-0.570	-0.570	0.000	-0.048	-0.618
						0.352	0.352	0.000	0.012	0.364
						0.423	0.423	0.000	0.070	0.493
						0.423	0.423	0.000	0.063	0.486
						0.200	0.200	0.000	0.010	0.210

Table A2-2. LFMM parameter values for platinum complexes. The atom types employed are the same as AMBER94.

Morse								
		r_0 (Å)	D_0 (kcal/mol)	α (1/Å)				
Pt+2	CL-	2.49079	52	1.44416				
Pt+2	HYDR	1.67686	75	2.08984				
Pt+2	N3	2.31653	46	1.63273				
Pt+2	BR-	2.55103	52	1.5331				
Pt+2	P	2.34522	53	1.94219				
Pt+2	OH2	2.33859	57	1.16197				
Pt+2	S	2.51364	52	1.54781				
Pt+2	NC	2.26093	60	1.67828				
Pt+2	NB	2.24093	60	1.72828				
Pt+2	C	2.05	75	2.0898				
A_{LL}/n		A_{LL} (kcal/mol)	n					
Pt+2	HYDR	247	2					
Pt+2	CL-	1850	4					
Pt+2	N3	2800	6					
Pt+2	BR-	2580	4					
Pt+2	P	4155	4					
Pt+2	OH2	2800	6					
Pt+2	S	770	4					
Pt+2	NC	2400	6					
Pt+2	NB	2400	6					
Pt+2	C	247	2					

[esig]						1/cm	1/cm	1/cm
Pt+2	CL-	0	0	0	0	0	0	1.65E+06
Pt+2	HYDR	0	0	0	0	0	0	147000
Pt+2	N3	0	0	0	0	0	610000	0
Pt+2	BR-	0	0	0	0	0	0	1.66E+06
Pt+2	P	0	0	0	0	380000	0	0
Pt+2	OH2	0	0	0	0	0	400000	0
Pt+2	S	0	0	0	0	0	1.02E+06	0
Pt+2	NC	0	0	0	0	0	0	1.33E+06
Pt+2	NB	0	0	0	0	0	0	1.33E+06
Pt+2	C	0	0	0	0	0	0	147000
[epix]						1/cm	1/cm	
Pt+2	CL-	0	0	0	0	0	175000	0
Pt+2	HYDR	0	0	0	0	0	0	0
Pt+2	N3	0	0	0	0	0	0	0
Pt+2	BR-	0	0	0	0	0	80000	0
Pt+2	P	0	0	0	0	-92000	0	0
Pt+2	OH2	0	0	0	0	0	0	0
Pt+2	S	0	0	0	0	0	0	0
Pt+2	NC	0	0	0	0	0	0	0
Pt+2	NB	0	0	0	0	0	0	0
Pt+2	C	0	0	0	0	0	0	0
[epiy]						1/cm	1/cm	1/cm

Pt+2	CL-	0	0	0	0	0	175000	0
Pt+2	HYDR	0	0	0	0	0	0	0
Pt+2	N3	0	0	0	0	0	0	0
Pt+2	BR-	0	0	0	0	0	80000	0
Pt+2	P	0	0	0	0	-92000	0	0
Pt+2	OH2	0	0	0	0	0	190000	0
Pt+2	S	0	0	0	0	93115	0	0
Pt+2	NC	0	0	0	0	0	0	590000
Pt+2	NB	0	0	0	0	0	0	590000
Pt+2	C	0	0	0	0	0	0	0
[exds]						1/cm	1/cm	1/cm
Pt+2	CL-	0	0	0	0	0	0	445000
Pt+2	HYDR	0	0	0	0	0	0	155000
Pt+2	N3	0	0	0	0	0	121000	0
Pt+2	BR-	0	0	0	0	96000	0	0
Pt+2	P	0	0	0	0	150000	0	0
Pt+2	OH2	0	0	0	0	0	0	163000
Pt+2	S	0	0	0	0	91000	0	0
Pt+2	NC	0	0	0	0	0	135000	0
Pt+2	NB	0	0	0	0	0	135000	0
Pt+2	C	0	0	0	0	0	0	155000

Table A2-3. LFMM-Optimized and QM/MM[a] M–L Bond Lengths (in Å) for double-stranded Pt Complexes.

	LFMM				QM/MM			
	Pt–N7A	Pt–N7B	Pt–N1	Pt–N2	Pt–N7A	Pt–N7B	Pt–N1	Pt–N2
1	2.037	2.040	2.014	2.005	2.019	2.016	2.035	2.045
2	2.008	2.000	2.030	2.032	2.014	2.024	2.027	2.039
3	2.012	1.998	2.030	2.027	2.026	2.016	2.023	2.036
4	2.011	1.999	2.031	2.029	2.016	2.025	2.038	2.027
5	2.012	2.001	2.030	2.040	2.018	2.029	2.037	2.033
6	2.015	2.003	2.030	2.053	2.025	2.026	2.024	2.047
7	2.014	2.002	2.042	2.034	2.021	2.030	2.032	2.038
8	2.023	2.006	2.052	2.064	2.031	2.019	2.037	2.065
9	2.032	2.011	2.063	2.070	2.031	2.022	2.047	2.064
10	2.011	2.000	2.027	2.033	2.020	2.025	2.024	2.035
11	2.011	1.999	2.028	2.031	2.017	2.026	2.026	2.032
12	2.012	1.998	2.025	2.029	2.019	2.028	2.019	2.031
13	2.011	1.999	2.029	2.034	2.020	2.025	2.028	2.036

[a] Gkionis, K.; Platts, J. A., *J. Biol. Inorg. Chem.* **2009,** *14* (8), 1165.

Table A2-4. The computed low energy LFMM and QM frequencies (cm^{-1}) for *cis*-$[Pt(NH_3)_2(Cl)(dGuo)]^+$ in Figure 4-1.

ADF	LFMM
16	23
35	35
43	41
45	62
65	74
84	89
95	115
114	134
118	137
125	160
134	164
155	178
161	189
177	195
195	218
207	223
211	250
222	261
235	282
245	286
251	307
275	317
292	327
308	345
314	352
324	360
336	367
339	377
370	386
394	418
422	451
429	484
440	499
455	544
457	561
470	572
506	579
572	585
592	601

Appendix 3

Table A3-1. Molecular dynamics simulation protocol used in MOE simulations.

1A84						
1782	H2O					
22	Na+					
md step	sub step	time step/ps	time/ps	ensemble	thermostat	constraint
emin	1					pt/DNA
heating	1		10	NVT, 0K to 15	Berendsen	pt/DNA
	2		10	NVT, 150K to	Berendsen	pt/DNA
equilibration	1	0.002	100	NVT, 300K	NHA	pt/DNA
opt	1					none
production		0.002	5000	NVT, 300K	NHA	none
1A2E						
1532	H2O					
18	Na+					
md step	sub step	time step/ps	time/ps	ensemble	thermostat	constraint
emin	1					pt/DNA
heating	1		10	NVT, 0K to 15	Berendsen	pt/DNA
	2		10	NVT, 150K to	Berendsen	pt/DNA
equilibration	1	0.002	100	NVT, 300K	NHA	pt/DNA
opt	1					none
production		0.002	5000	NVT, 300K	NHA	none

Table A3-2. Molecular dynamics simulation protocol used in DL_POLY simulations.

1A2E											
General settings:											
cut	rvdW	delR	shake tol	ewald precision		no link	Berendsen NVT or NPT		Verlet leap-frog		
12.0 A	12.0 A	0.5 A	1.00E-08	1.00E-06							
Sim	ensemble	T /K	p / atm	tau T / ps	tau p / ps	timestep / ps	steps	equilibration	scale	only water bonds frozen	
1	NPT	300	1	0.01	0.1	0.0005	50000	50000	1	+	
2	NPT	300	1	0.1	0.3	0.001	10000	10000	1	+	
3	NPT	300	1	0.3	0.4	0.001	10000	10000	1	+	
4	NPT	300	1	1	1	0.001	10000	10000	1	+	
5	NPT	300	1	2	2	0.001	5000	5000	1	+	
6	NPT	300	1	0.2	0.2	0.002	50000	50000	1	+	
7	NPT	300	1	0.5	0.5	0.002	50000	0	1	+	
8	NPT	300	1	1	1	0.002	50000	0	1	+	
9	NPT	300	1	2	2	0.002	50000	0	1	+	
10	NPT	300	1	2	2	0.002	3000000	0	1	+	
2	Cl-										
18	Na+										
3941	H2O										

1A2E DE-PLATINATED											
General settings:											
cut	rvdW	delR	shake tol	ewald precision		no link	Berendsen NVT or NPT		Verlet leap-frog		
12.0 A	12.0 A	0.5 A	1.00E-08	1.00E-06							
Sim	ensemble	T /K	p / atm	tau T / ps	tau p / ps	timestep / ps	steps	equilibration	scale	only water bonds frozen	
1	NPT	300	1	0.1	0.5	0.0001	50000	50000	1	+	
2	NPT	300	1	0.1	0.2	0.0005	5000	5000	1	+	
3	NPT	300	1	0.1	0.1	0.001	5000	5000	1	+	
4	NPT	300	1	0.5	0.1	0.001	50000	50000	1	+	
5	NPT	300	1	1	1	0.001	50000	50000	1	+	
6	NPT	300	1	2	2	0.001	50000	50000	1	+	
7	NPT	300	1	2	2	0.002	3000000	50000	1	+	
18	Na+										
3941	H2O										

P-DNA													
General settings:													
cut	rvdW	delR	shake tol	ewald precision		no link	Berendsen NVT or NPT		velocity Verlet				
12.0 A	12.0 A	0.5 A	1.00E-08	1.00E-06									
Sim	ensemble	T /K	p / atm	tau T / ps	tau p / ps	timestep / ps	steps	equilibration	scale	only water bonds frozen			
1	NPT	300	1	0.3	0.3	0.0005	50000	50000	1	+			
2	NPT	300	1	0.3	0.3	0.001	50000	50000	1	+			
3	NPT	300	1	0.5	0.5	0.002	100000	100000	1	+			
4	NPT	300	1	0.5	0.5	0.002	100000	0	0	+			
5	NPT	300	1	0.2	0.2	0.002	500000	500000	1	+			
6	NVT	450		0.2		0.003	5000	5000	1	---			
7	NVT	300		0.2		0.002	50000	50000	15	+			
8	NVT	450		0.2		0.003	5000	5000	1	---			
9	NVT	300		0.2		0.002	50000	50000	15	+			
10	NVT	450		0.2		0.003	5000	5000	1	---			
11	NVT	300		0.2		0.002	50000	50000	15	+			
12	NPT	300	1	0.5	0.5	0.002	250000	250000	1	+			
13	NPT	300	1	0.5	0.5	0.002	250000	250000	1	+			
14	NPT	300	1	1	1	0.002	250000	250000	25	+			
15	NPT	300	1	2	2	0.002	250000	250000	25	+			
16	NPT	300	1	2	2	0.002	250000	250000	25	+			
17	NPT	300	1	2	2	0.002	250000	250000	100	+			
18	NPT	300	1	2	2	0.002	3000000	3000000	100	+			
22	Na+	2	Cl-							---	means all bonds except (Pt-L) bonds		
3400	H2O										were frozen to reference value		

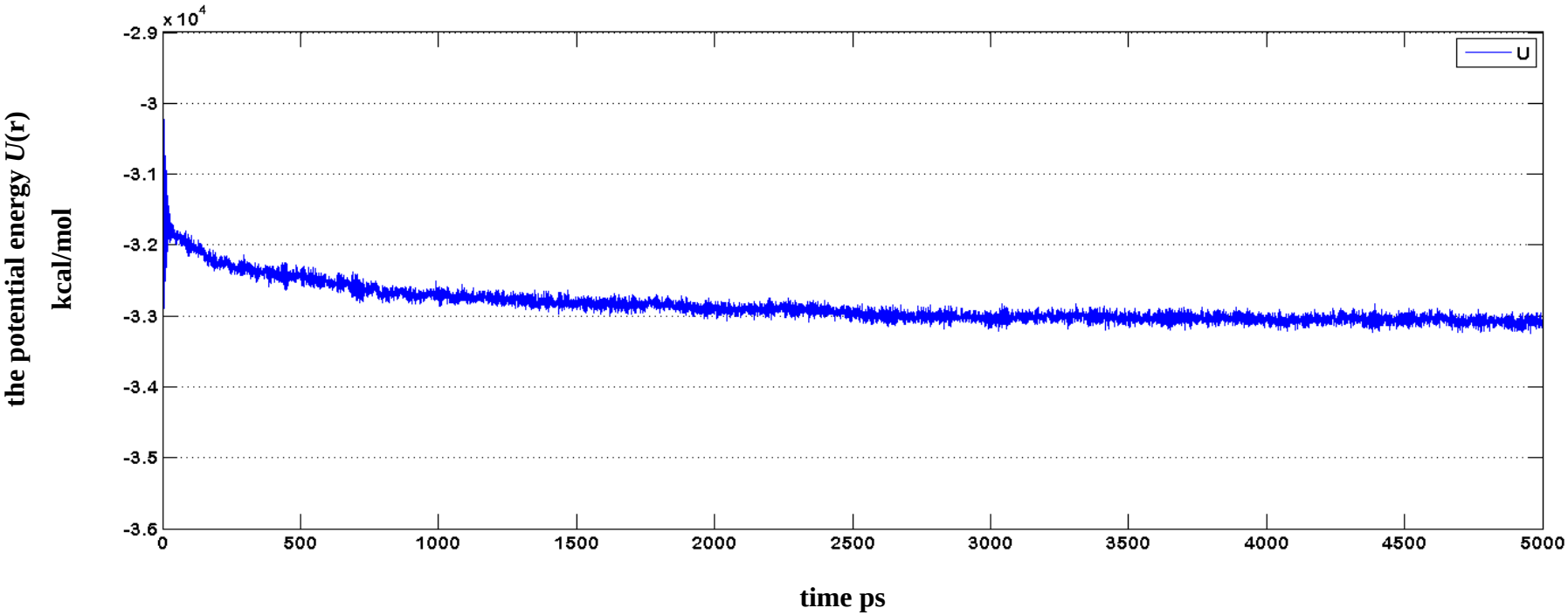


Figure A3-1. The potential energy $U(\mathbf{r})$ of 1A84 at time t in kcal/mol during LFMD simulations using DommiMOE.

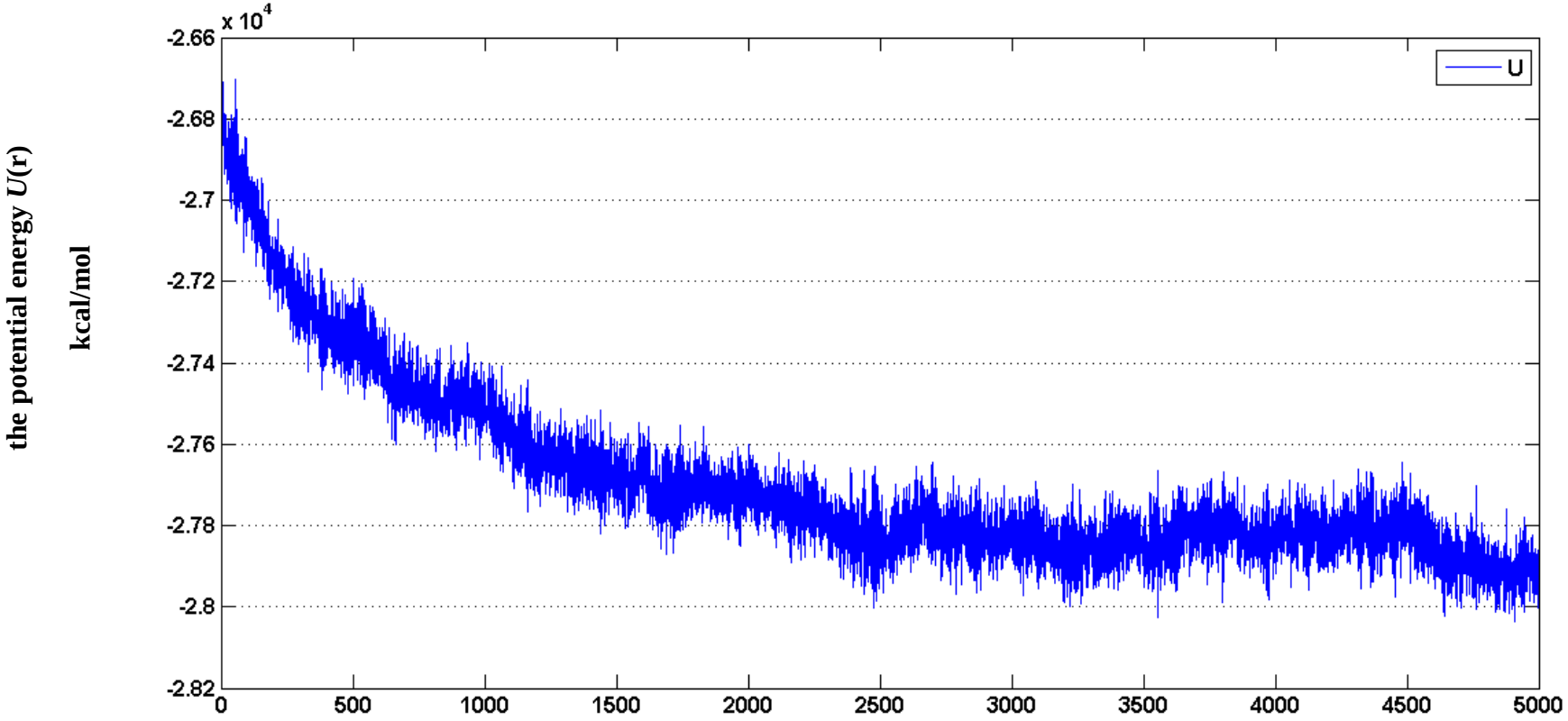


Figure A3-2. The potential energy $U(\mathbf{r})$ of 1A2E at time t in kcal/mol during LFMD simulations using DommiMOE.

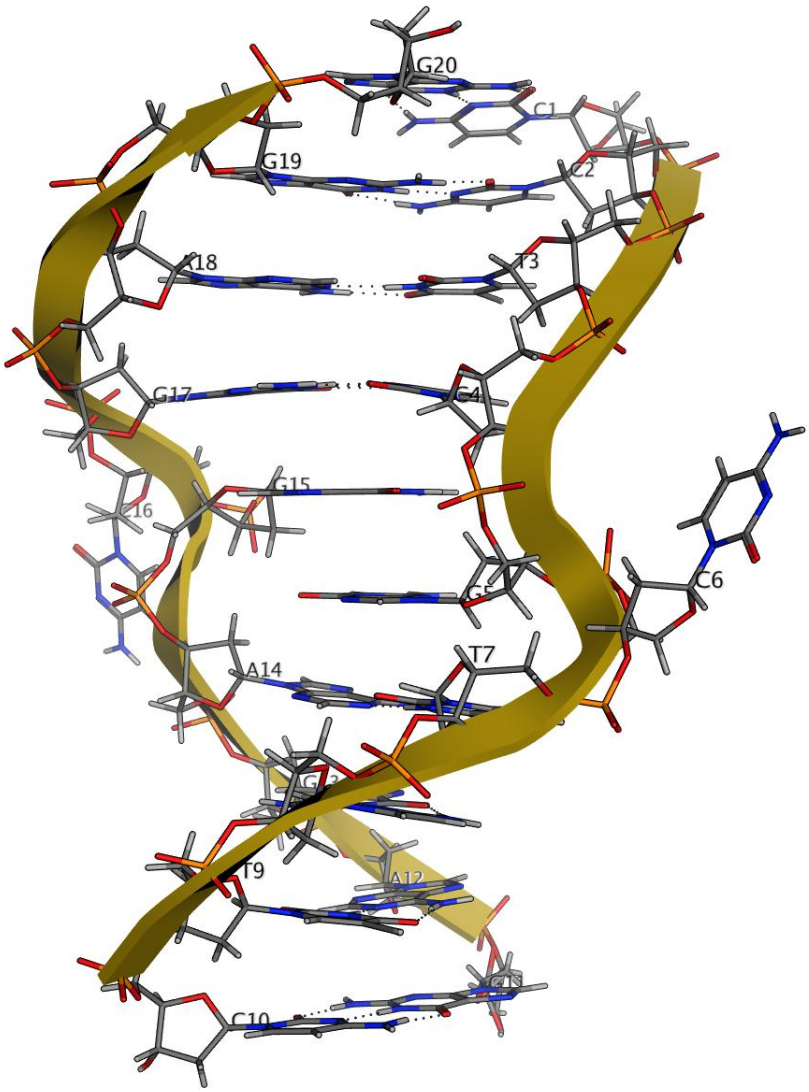


Figure A3-3. The average LFMD structure derived from simulations of deplatinated 1A2E with periodic boundary conditions.

References

(1) Kelland, L. *Nat. Rev. Cancer* **2007**, *7*, 573.

(2) Lippert, B. *Chemistry and biochemistry of a leading anticancer drug*; Wiley-VCH: Weinheim, 1999.

(3) Rosenberg, B.; Van Camp, L.; Krigas, T. *Nature* **1965**, *205*, 698.

(4) Weiss, R.; Christian, M. *Drugs* **1993**, *46*, 360.

(5) Wong, E.; Giandomenico, C. M. *Chem. Rev.* **1999**, *99*, 2451.

(6) Jung, Y.; Lippard, S. J. *Chem. Rev.* **2007**, *107*, 1387.

(7) Fuertes, M. A.; Alonso, C.; Pérez, J. M. *Chem. Rev.* **2003**, *103*, 645.

(8) Bradley, L. J. N.; Yarema, K. J.; Lippard, S. J.; Essigmann, J. M. *Biochemistry* **1993**, *32*, 982.

(9) Coste, F.; Malinge, J.; Serre, L.; Shepard, W.; Roth, M.; Leng, M.; Zelwer, C. *Nucleic Acids Res.* **1999**, *27*, 1837.

(10) Gelasco, A.; Lippard, S. J. *Biochemistry* **1998**, *37*, 9230.

(11) Jamieson, E. R.; Lippard, S. J. *Chem. Rev.* **1999**, *99*, 2467.

(12) Elizondo-Riojas, M. A.; Kozelka, J. *J. Mol. Biol.* **2001**, *314*, 1227.

(13) Zhu, Y. Y.; Wang, Y.; Chen, G. J. *Nucleic Acids Res.* **2009**, *37*, 5930.

(14) Teletchea, S.; Skauge, T.; Sletten, E.; Kozelka, J. *Chem. Eur. J.* **2009**, *15*, 12320.

(15) Barnham, K. J.; Berners-Price, S. J.; Frenkiel, T. A.; Frey, U.; Sadler, P. J. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1874.

(16) Parkinson, J. A.; Chen, Y.; Murdoch, P. D.; Guo, Z. J.; Berners-Price, S. J.; Brown, T.; Sadler, P. J. *Chem. Eur. J.* **2000**, *6*, 3636.

(17) Chen, Y.; Parkinson, J. A.; Guo, Z. J.; Brown, T.; Sadler, P. J. *Angew. Chem. Int. Ed.* **1999**, *38*, 2060.

(18) Reedijk, J. *Platinum Metals Review* **2008**, *52*, 2.

(19) *Cancer Drug Resistance*; Teicher, B. A., Ed.; Humana Press: Totowa, 2006.

(20) *Studies on the antitumor activity of group VIII transition metal complexes. Part I. Platinum (II) complexes*; Cleare, M. J.; Hoeschele, J. D., Eds., 1973; Vol. 2.

(21) *Metal Ions in Biological Systems: Metal Complexes in Tumor Diagnosis and as Anticancer Agents*; Sigel, A.; Sigel, H., Eds.; Marcel Dekker: New York, 2004; Vol. 42.

(22) White, S. C.; Lorigan, P.; Margison, G. P.; Margison, J. M.; Martin, F.; Thatcher, N.; Anderson, H.; Ranson, M. *Br. J. Cancer* **2006**, *95*, 822.

(23) Dragovich, T.; Mendelson, D.; Kurtin, S.; Richardson, K.; Von Hoff, D.; Hoos, A. *Cancer Chemother. Pharmacol.* **2006**, *58*, 759.

(24) Rademaker-Lakhai, J. M.; Terret, C.; Howell, S. B.; Baud, C. M.; de Boer, R. F.; Pluim, D.; Beijnen, J. H.; Schellens, J. H. M.; Droz, J. P. *Clin. Cancer Res.* **2004**, *10*, 3386.

(25) Bednarski, P. J.; Mackay, F. S.; Sadler, P. J. *Anti-Cancer Agents Med. Chem.* **2007**, *7*, 75.

(26) Brown, S. B.; Brown, E. A.; Walker, I. *Lancet Oncol.* **2004**, *5*, 497.

(27) Wayne, C. E.; Wayne, R. P. *Photochemistry*; 1st ed.; Oxford University Press: New York, 1996.

(28) Atkins, P. W.; De Paula, J. *Atkins' Physical chemistry*; 8th ed.; Oxford University Press: Oxford, 2006.

(29) Farrer, N. J.; Salassa, L.; Sadler, P. J. *Dalton Trans.* **2009**, 10690.

(30) *Photochemistry and Photophysics of Coordination Compounds I*; Balzani, V.; Campagna, S., Eds.; Springer: Berlin, 2007; Vol. 280.

(31) van Rijt, S. H.; Sadler, P. J. *Drug Discov. Today* **2009**, *14*, 1089.

(32) Bruijnincx, P. C. A.; Sadler, P. J. *Adv. Inorg. Chem.* **2009**, *61*, 1.

(33) Bednarski, P. J.; Grunert, R.; Zielzki, M.; Wellner, A.; Mackay, F. S.; Sadler, P. J. *Chem. Biol.* **2006**, *13*, 61.

(34) Mackay, F. S.; Woods, J. A.; Moseley, H.; Ferguson, J.; Dawson, A.; Parsons, S.; Sadler, P. J. *Chem. Eur. J.* **2006**, *12*, 3155.

(35) Mackay, F. S.; Woods, J. A.; Heringova, P.; Kasparkova, J.; Pizarro, A. M.; Moggach, S. A.; Parsons, S.; Brabec, V.; Sadler, P. J. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20743.

(36) Farrer, N. J.; Woods, J. A.; Salassa, L.; Zhao, Y.; Robinson, K. S.; Clarkson, G.; Mackay, F. S.; Sadler, P. J. *Angew. Chem. Int. Ed.* **2010**, *49*, 8905.

(37) Kozelka, J. *Inorg. Chim. Acta* **2009**, *362*, 651.

(38) Butour, J. L.; Johnson, N. P. *Biochemistry* **1986**, *25*, 4534.

(39) Arpalahti, J.; Lippert, B. *Inorg. Chem.* **1990**, *29*, 104.

(40) Baik, M.-H.; Friesner, R. A.; Lippard, S. J. *J. Am. Chem. Soc.* **2003**, *125*, 14082.

(41) Raber, J.; Zhu, C.; Eriksson, L. A. *J. Phys. Chem. B* **2005**, *109*, 11006.

(42) Monjardet-Bas, V.; Elizondo-Riojas, M.-A.; Chottard, J.-C.; Kozelka, J. *Angew. Chem. Int. Ed.* **2002**, *41*, 2998.

(43) Laoui, A.; Kozelka, J.; Chottard, J. C. *Inorg. Chem.* **1988**, *27*, 2751.

(44) Legendre, F.; Bas, V.; Kozelka, J.; Chottard, J.-C. *Chem. Eur. J.* **2000**, *6*, 2002.

(45) Monjardet-Bas, V.; Chottard, J.-C.; Kozelka, J. *Chem. Eur. J.* **2002**, *8*, 1144.

(46) Monjardet-Bas, V.; Bombard, S.; Chottard, J.-C.; Kozelka, J. *Chem. Eur. J.* **2003**, *9*, 4739.

(47) Reeder, F.; Gonnet, F.; Kozelka, J.; Chottard, J.-C. *Chem. Eur. J.* **1996**, *2*, 1068.

(48) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482.

(49) Huppert, J. L. *Chem. Soc. Rev.* **2008**, *37*, 1375.

(50) Hausheer, F. H.; Pavankumar, P. N. V.; Seetharamulu, P.; Yao, S.; Saxe, J. D.; Reddy, D. G. *J. Comput. Chem.* **1999**, *20*, 365.

(51) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876.

(52) Read, M.; Neidle, S. *Biochemistry* **2000**, *39*, 13422.

(53) Tauchi, T.; Shin-ya, K.; Sashida, G.; Sumi, M.; Okabe, S.; Ohyashiki, J. H.; Ohyashiki, K. *Oncogene* **2006**, *25*, 5719.

(54) Kieltyka, R.; Englebienne, P.; Fakhoury, J.; Autexier, C.; Moitessier, N.; Sleiman, H. F. *J. Am. Chem. Soc.* **2008**, *130*, 10040.

(55) Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345.

(56) Sparapani, S.; Haider, S. M.; Doria, F.; Gunaratnam, M.; Neidle, S. *J. Am. Chem. Soc.* **2010**, *132*, 12263.

(57) E.J. Baerends, J. A., A. Bérces, F.M. Bickelhaupt, C. Bo, P.M. Boerrigter, L. Cavallo, D.P. Chong, L. Deng, R.M. Dickson, D.E. Ellis, M. van Faassen, L. Fan, T.H. Fischer, C. Fonseca Guerra, S.J.A. van Gisbergen, J.A. Groeneveld, O.V. Gritsenko, M. Grüning, F.E. Harris, P. van den Hoek, C.R. Jacob, H. Jacobsen, L. Jensen, G. van Kessel, F. Kootstra, E. van Lenthe, D.A. McCormack, A. Michalak, J. Neugebauer, V.P. Nicu, V.P. Osinga, S. Patchkovskii, P.H.T. Philipsen, D. Post, C.C. Pye, W. Ravenek, P. Ros, P.R.T. Schipper, G.

Schreckenbach, J.G. Snijders, M. Solà, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T.A. Wesolowski, E.M. van Wezenbeek, G. Wiesenekker, S.K. Wolff, T.K. Woo, A.L. Yakovlev, and T. Ziegler; ADF2007.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2007.

(58) van Lenthe, E.; van Leeuwen, R.; Baerends, E. J.; Snijders, J. G. *Int. J. Quantum Chem.* **1996**, *57*, 281.

(59) van Lenthe, E.; Ehlers, A.; Baerends, E.-J. *J. Chem. Phys.* **1999**, *110*, 8943.

(60) Schipper, P. R. T.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2000**, *112*, 1344.

(61) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Montgomery Jr, J.; Vreven, T.; Kudin, K.; Burant, J. In *Gaussian 2003*; Revision B. 05, Gaussian, Inc.: Pittsburgh, PA, 2003.

(62) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(63) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(64) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270.

(65) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

(66) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669.

(67) Cossi, M.; Barone, V. *J. Chem. Phys.* **2001**, *115*, 4708.

(68) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.

(69) Mackay, F. S.; Farrer, N. J.; Salassa, L.; Tai, H. C.; Deeth, R. J.; Moggach, S. A.; Wood, P. A.; Parsons, S.; Sadler, P. J. *Dalton Trans.* **2009**, 2315.

(70) Hambley, T. W.; Jones, A. R. *Coord. Chem. Rev.* **2001**, *212*, 35.

(71) Chval, Z.; Sip, M. *J. Phys. Chem. B* **1998**, *102*, 1659.

(72) Deeth, R. J.; Anastasi, A.; Diedrich, C.; Randell, K. *Coord. Chem. Rev.* **2009**, *253*, 795.

(73) Cramer, C. J. *Essentials of computational chemistry: theories and models*; John Wiley & Sons Inc: Chichester, 2004.

(74) Gkionis, K.; Platts, J. A. *J. Biol. Inorg. Chem.* **2009**, *14*, 1165.

(75) Spiegel, K.; Rothlisberger, U.; Carloni, P. *J. Phys. Chem. B* **2004**, *108*, 2699.

(76) Carloni, P.; Rothlisberger, U.; Parrinello, M. *Acc. Chem. Res* **2002**, *35*, 455.

(77) Deeth, R. J.; Diedrich, C. *J. Biol. Inorg. Chem.* **2010**, *15*, 117.

(78) Cheatham, T. E., III.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845.

(79) Takahara, P. M.; Rosenzweig, A. C.; Frederick, C. A.; Lippard, S. J. *Nature* **1995**, *377*, 649.

(80) Huang, H.; Zhu, L.; Reid, B.; Drobny, G.; Hopkins, P. *Science* **1995**, *270*, 1842.

(81) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63.

(82) Rosa, A.; Ricciardi, G.; Gritsenko, O.; Baerends, E. *Struct. Bonding (Berlin)* **2004**, *112*, 49.

(83) Marques, M. A. L.; Gross, E. K. U. *Annu. Rev. Phys. Chem.* **2004**, *55*, 427.

(84) Chong, D. P. *Recent advances in density functional methods*; World Scientific: Singapore, 1995; Vol. 1.

(85) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997.

(86) Gross, E. K. U.; Kohn, W. *Adv. Quant. Chem.* **1990**, *21*, 255.

(87) Gruning, M.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2002**, *116*, 9591.

(88) Gritsenko, O.; van Leeuwen, R.; van Lenthe, E.; Baerends, E. J. *Phys. Rev. A* **1995**, *51*, 1944.

(89) Leach, A. R. *Molecular modelling: principles and applications*; Pearson Education: Harlow 2001.

(90) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(91) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(92) MacKerell, A. D.; Wiorkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946.

(93) Foloppe, N.; MacKerell Jr, A. *J. Comput. Chem.* **2000**, *21*, 86.

(94) MacKerell Jr, A.; Banavali, N. *J. Comput. Chem.* **2000**, *21*, 105.

(95) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham Iii, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817.

(96) Cheatham III, T.; Young, M. *Biopolymers* **2001**, *56*, 232.

(97) Ponomarev, S.; Thayer, K.; Beveridge, D. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14771.

(98) Burton, V. J.; Deeth, R. J.; Kemp, C. M.; Gilbert, P. J. *J. Am. Chem. Soc.* **1995**, *117*, 8407.

(99) Deeth, R. J. In *Computational Inorganic and Bioinorganic Chemistry*; 3rd ed.; Solomon, E. I., King, R. B., Scott, R. A., Eds.; John Wiley & Sons: Chichester, 2009.

(100) Deeth, R. J. *Coord. Chem. Rev.* **2001**, *212*, 11.

(101) Deeth, R. J.; Fey, N.; Williams-Hubbard, B. *J. Comp. Chem.* **2005**, *26*, 123.

(102) Deeth, R. J.; Anastasi, A. E.; Wilcockson, M. J. *J. Am. Chem. Soc.* **2010**, *132*, 6876.

(103) Huheey, J. E.; Keiter, E. A.; Keiter, R. L. *Inorganic chemistry: principles of structure and reactivity*; 4th ed.; Harper Collins: New York, 1993.

(104) Figgis, B. N.; Hitchman, M. A. *Ligand field theory and its applications*; Wiley-VCH: New York, 2000.

(105) A copy of DommiMOE can be obtained from the authors. Certain conditions apply.

(106) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.

(107) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520.

(108) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553.

(109) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587.

(110) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616.

(111) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.

(112) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(113) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press: Oxford, 1990.

(114) *Computational Biochemistry and Biophysics*; Becker, O. M.; MacKerell Jr., A. D.; Roux, B.; Watanabe, M., Eds.; Marcel Dekker: New York, 2001.

(115) Verlet, L. *Phys. Rev.* **1967**, *159*, 98.

(116) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637.

(117) MOE. Chemical Computing Group: Montreal, 2009.

(118) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(119) Ewald, P. *Annalen der Physik* **1921**, *64*, 253.

(120) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(121) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384.

(122) Nosé, S. *Mol. Phys.* **1984**, *53*, 255.

(123) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

(124) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931.

(125) Morokuma, K. *J. Chem. Phys.* **1971**, *55*, 1236.

(126) Ziegler, T.; Rauk, A. *Theor. Chem. Acc.* **1977**, *46*, 1.

(127) *Reviews of Computational Chemistry; Boyd, DB; Lipkowitz, KB, Eds.*; Bickelhaupt, F. M.; Baerends, E. J., Eds.; Wiley-VCH: New York, 2000; Vol. 15.

(128) Frenking, G.; Frohlich, N. *Chem. Rev.* **2000**, *100*, 717.

(129) Wang, F.; Ziegler, T. *J. Chem. Phys.* **2005**, *123*, 194102.

(130) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098.

(131) Perdew, J. *Phys. Rev. B* **1986**, *33*, 8822.

(132) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.

(133) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.

(134) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.

(135) Ehlers, A.; Böhme, M.; Dapprich, S.; Gobbi, A.; Höllwarth, A.; Jonas, V.; Köhler, K.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, *208*, 111.

(136) IUPAC-IUB Joint Commission on Biochemical Nomenclature *Eur. J. Biochem.* **1983**, *131*, 5.

(137) IUPAC-IUB Joint Commission on Biochemical Nomenclature *Eur. J. Biochem.* **1983**, *131*, 9.

(138) Lavery, R.; Moakher, M.; Maddocks, J.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 5917.

(139) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108.

(140) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry. International Edition*; 5th ed.; W. H. Freeman and Company: New York, 2002.

(141) Rodger, A.; Nordén, B. *Circular dichroism and linear dichroism*; Oxford University Press: Oxford, 1996.

(142) Harrap, K. *Cancer Treat Rev* **1985**, *12*, 21.

(143) Knox, R. J.; Friedlos, F.; Lydall, D. A.; Roberts, J. J. *Cancer Res.* **1986**, *46*, 1972.

(144) Kelland, L. *Cancer Res.* **1993**, *53*, 2581.

(145) Kelland, L. *Exp. Opin. Invest. Drugs* **2000**, *9*, 1373.

(146) Kelland, L. R. *Cancer Res.* **1992**, *52*, 3857.

(147) Gately, D. P.; Howell, S. B. *Br. J. Cancer* **1993**, *67*, 1171.

(148) Ishida, S.; Lee, J.; Thiele, D. J.; Herskowitz, I. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 14298.

(149) Mistry, P.; Kelland, L. R.; Abel, G.; Sidhar, S.; Harrap, K. R. *Br. J. Cancer* **1991**, *64*, 215.

(150) Dabholkar, M.; Bostick-Bruton, F.; Weber, C.; Bohr, V. A.; Egwuagu, C.; Reed, E. *J. Natl. Cancer Inst.* **1992**, *84*, 1512.

(151) Gifford, G.; Paul, J.; Vasey, P.; Kaye, S.; Brown, R. *Clin. Cancer Res.* **2004**, *10*, 4420.

(152) Kratz, F.; Mueller, I.; Ryppa, C.; Warnecke, A. *ChemMedChem* **2008**, *3*, 20.

(153) Farrer, N. J.; Sadler, P. J. *Aust. J. Chem.* **2008**, *61*, 669.

(154) Kasparkova, J.; Heringova, P.; Brabec, V.; Mackay, F.; Sadler, P. J. *Metal Ions in Biology and Medicine, Vol 10* **2008**, *10*, 463.

(155) Kasparkova, J.; Mackay, F. S.; Brabec, V.; Sadler, P. J. *J. Biol. Inorg. Chem.* **2003**, *8*, 741.

(156) Kratochwil, N. A.; Ivanov, A. I.; Patriarca, M.; Parkinson, J. A.; Gouldsworthy, A. M.; Murdoch, P. D.; Sadler, P. J. *J. Am. Chem. Soc.* **1999**, *121*, 8193.

(157) Mackay, F. S.; Woods, J.; Moseley, H.; Sadler, P. J. *Br. J. Dermatol.* **2005**, *152*, 857.

(158) Muller, P.; Schroder, B.; Parkinson, J. A.; Kratochwil, N. A.; Coxall, R. A.; Parkin, A.; Parsons, S.; Sadler, P. J. *Angew. Chem. Int. Ed.* **2003**, *42*, 335.

(159) Phillips, H. I. A.; Ronconi, L.; Sadler, P. J. *Chem. Eur. J.* **2009**, *15*, 1588.

(160) Brancaleon, L.; Moseley, H. *Lasers Med. Sci.* **2002**, *17*, 173.

(161) Sokolov, A. Y.; Schaefer III, H. F. *Dalton Trans.* **2011**, *40*, 7571.

(162) Hall, M.; Hambley, T. *Coord. Chem. Rev.* **2002**, *232*, 49.

(163) Hall, M. D.; Mellor, H. R.; Callaghan, R.; Hambley, T. W. *J. Med. Chem.* **2007**, *50*, 3403.

(164) Mackay, F. S.; Moggach, S. A.; Collins, A.; Parsons, S.; Sadler, P. J. *Inorg. Chim. Acta* **2009**, *362*, 811.

(165) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.

(166) Petzold, H.; Xu, J. J.; Sadler, P. J. *Angew. Chem. Int. Ed.* **2008**, *47*, 3008.

(167) Detty, M. R.; Gibson, S. L.; Wagner, S. J. *J. Med. Chem.* **2004**, *47*, 3897.

(168) Loganathan, D.; Morrison, H. *Photochem. Photobiol.* **2006**, *82*, 237.

(169) Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224.

(170) Butler, R. S.; Cohn, P.; Tenzel, P.; Abboud, K. A.; Castellano, R. K. *J. Am. Chem. Soc.* **2009**, *131*, 623.

(171) Wagenknecht, P. S.; Ford, P. C. *Coord. Chem. Rev.* **2011**, *255*, 591.

(172) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.

(173) Anastasi, A.; Deeth, R. J. *J. Chem. Theory Comput.* **2009**, *5*, 2339.

(174) Allen, F. *Acta Crystallogr. Sect. B: Struct. Sci.* **2002**, *58*, 380.

(175) Sherman, S.; Gibson, D.; Wang, A.; Lippard, S. *Science* **1985**, *230*, 412.

(176) Coll, M.; Sherman, S.; Gibson, D.; Lippard, S.; Wang, A. *J. Biomol. Struct. Dyn.* **1990**, *8*, 315.

(177) Noll, D. M.; Mason, T. M.; Miller, P. S. *Chem. Rev.* **2006**, *106*, 277.

(178) Natile, G.; Coluccia, M. *Coord. Chem. Rev.* **2001**, *216*, 383.

(179) Wu, B.; Davey, C. A. *Chem. Biol.* **2008**, *15*, 1023.

(180) Ober, M.; Lippard, S. J. *J. Am. Chem. Soc.* **2008**, *130*, 2851.

(181) Ober, M.; Lippard, S. J. *J. Am. Chem. Soc.* **2007**, *129*, 6278.

(182) Galea, A. M.; Murray, V. *Chem. Biol. Drug Des.* **2010**, *75*, 578.

(183) Smith, W.; Yong, C. W.; Rodger, P. M. *Mol. Simul.* **2002**, *28*, 385.

(184) Lovejoy, K. S.; Todd, R. C.; Zhang, S. Z.; McCormick, M. S.; D'Aquino, J. A.; Reardon, J. T.; Sancar, A.; Giacomini, K. M.; Lippard, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8902.

(185) Kartalou, M.; Essigmann, J. M. *Mutat. Res.* **2001**, *478*, 1.

(186) Rohs, R.; West, S. M.; Sosinsky, A.; Liu, P.; Mann, R. S.; Honig, B. *Nature* **2009**, *461*, 1248.

(187) Nikolov, D.; Chen, H.; Halay, E.; Hoffman, A.; Roeder, R.; Burley, S. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 4862.

(188) Weinberg, R. A. *The biology of cancer*; Garland Science: New York, 2007.

(189) Mantri, Y.; Lippard, S. J.; Baik, M. H. *J. Am. Chem. Soc.* **2007**, *129*, 5023.

(190) Deubel, D. V.; Lau, J. K. C. *Chem. Commun. (Cambridge, U. K.)* **2006**, 2451.

(191) Dougan, S. J.; Habtemariam, A.; McHale, S. E.; Parsons, S.; Sadler, P. J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 11628.

(192) Betanzos-Lara, S.; Salassa, L.; Habtemariam, A.; Sadler, P. J. *Chem. Commun. (Cambridge, U. K.)* **2009**, 6622.

(193) Ronconi, L.; Sadler, P. J. *Dalton Trans.* **2011**, *40*, 262.

(194) *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Chipot, C.; Pohorille, A., Eds.; Springer-Verlag: Berlin, 2007; Vol. 86.