Action of the anticancer drug cisplatin investigated by molecular simulation

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Trung Hai Nguyen, M.Sc.

aus An Giang, Vietnam

Berichter:
Universitätsprofessor Ph.D. Paolo Carloni
Professor Dr. Evert Jan Meijer

Tag der mündlichen Prüfung: 08.02.2013
Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

I hereby affirm that I composed this work independently and used no other than the specified sources and tools and that I marked all quotes as such.

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Aachen, September 24th, 2012

Trung Hai Nguyen

Bibliographic Information

Part of this thesis has been published at:

Nguyen, T. H.; Arnesano, F.; Scintilla, S.; Rossetti, G.; Ippoliti, E.; Carloni, P.; Natile, G. Structural Determinants of Cisplatin and Transplatin Binding to the Met-Rich Motif of Ctr1: A Computational Spectroscopy Approach. *Journal of Chemical Theory and Computation* **2012**, *8*, 2912–2920.

Abstract

Cancer is one of the leading causes of death worldwide. Among various chemotherapeutic drugs currently used in cancer treatment, cisplatin is one of the most important and successful drugs. It is the best hope for patients with testicular, ovarian and many other solid tumors. Cisplatin kills cancer cells by covalently binding to their DNA and causing DNA damages (cisplatin-DNA adducts). Albeit very efficient, cisplatin's clinical utility is limited due to its toxic side effects and resistance.

The mechanisms behind drug resistance are far from being completely understood. Experimental studies so far have suggested that binding of cisplatin to many cellular proteins is a key mechanism of drug resistance. In particular, the copper influx protein *Ctr1* has been shown to be responsible for the decrease of cellular uptake of cisplatin, commonly found in resistant cells. Furthermore, even if cisplatin can successfully reach and bind to the DNA, cancer cells may be able to repair the cisplatin-DNA adducts by employing a number of proteins involved in the so-called Nucleotide Excision Repair (NER) pathway.

This thesis has contributed to investigating processes, which are relevant to cisplatin resistance mechanism, using computational physics tools.

(i) First, we have studied binding of cisplatin and transplatin (a pharmacologically inactive *trans* isomer of cisplatin) to a model peptide (called *Mets7*) that mimics one of the methionine(Met)-rich motifs of *Ctr1*. We have used a set of computational tools including QM/MM simulations, force matching technique, replica exchange MD simulations and computational spectroscopy to predict the structures of the adducts formed due to the binding of cisplatin and transplatin to *Mets7*. Comparison between calculated and experimental spectroscopic data such as NMR, EXAFS and CD allows us to validate our predicted structures and to propose the most relevant species in water solution. The structural models arising from our calculations may serve as the molecular basis for further attempts to understand cisplatin uptake mediated by *Ctr1*.

(ii) Then, we have performed preliminary calculations on the interaction between cisplatin-DNA adduct and the High Mobility Group Box 1 (*HMGB1*) protein. *HMGB1* (a DNA-binding protein) has been shown to be able to block the NER repair pathway by binding to the cisplatin-DNA adducts and shielding them from being repaired by NER proteins. Although experimental studies have shown the structural determinants of the cisplatin-DNA-*HMGB1* complex and also provided the binding affinity, the mechanism of molecular recognition between cisplatin-DNA adduct and *HMGB1* is still unknown. We aim at investigating structural and energetic insights of the binding of *HMGB1* to the cisplatin-DNA adduct by applying free energy calculations based on metadynamics method. This thesis presents necessary calculations towards this goal.

Kurzfassung

Krebs ist eine der häufigsten Todesursachen weltweit. Unter den verschiedenen Chemotherapeutika, welche derzeit in der Krebsbehandlung eingesetzt werden, ist Cisplatin einer der wichtigsten und erfolgreichsten Wirkstoffe. Er ist die beste therapeutische Maßnahme für Patienten mit Hoden-, Eierstock- und vielen anderen soliden Tumoren. Cisplatin tötet Krebszellen durch kovalente Bindung an ihre DNA, wodurch Schäden an der DNA hervorgerufen werden (Cisplatin-DNA Addukte). Trotz sehr großer Effizienz ist die klinische Verwendbarkeit von Cisplatin aufgrund der toxischen Nebenwirkungen und Resistenzen beschränkt.

Die Mechanismen der Wirkstoff-Resistenzen sind bei weitem noch nicht vollständig verstanden. Experimentelle Studien haben bislang vorgeschlagen, dass die Bindung von Cisplatin an viele zelluläre Proteine ein wichtiger Resistenz-Mechanismus ist. Insbesondere wurde gezeigt, dass das Kupfer-Transportprotein *Ctr1* für die Verringerung der zellulären Aufnahme von Cisplatin, was in resistenten Zellen häufig gefunden wird, verantwortlich ist. Selbst wenn Cisplatin die DNA erfolgreich erreichen und anbinden kann, können Krebszellen in der Lage sein, die Cisplatin-DNA-Addukte unter Verwendung einer Reihe von Proteinen, welche in den sogenannten Nukleotidexzisionsreparatur-(NER)-Pfad eingebunden sind, zu reparieren.

Diese Doktorarbeit hat zur Untersuchung von Prozessen, welche für Cisplatin-Resistenz Mechanismen relevant sind, unter Verwendung von Computational Physics Methoden beigetragen.

(i) Zunächst haben wir die Bindung von Cisplatin und Transplatin (ein pharmakologisch inaktives trans-Isomer von Cisplatin) an ein Modell-Peptid (*Mets7*) untersucht, welches eines der Methionin(Met)-reichen Motive von *Ctr1* nachahmt. Wir haben eine Reihe von computergestützten Methoden einschließlich QM/MM-Simulationen, Force-Matching-Verfahren, Replica-Exchange MD Simulationen und computergestützte Spektroskopie verwendet, um Strukturen der Addukte, welche durch die Bindung von Cisplatin und Transplatin an *Mets7* gebildet werden, vorherzusagen. Ein Vergleich zwischen berechneten und experimentellen spektroskopischen Daten wie NMR, EXAFS und CD ermöglicht es uns, unsere vorhergesagten Strukturen zu validieren und die wichtigsten Spezies in wässriger Lösung vorzuschlagen. Die aus unseren Berechnungen hervorgegangenen strukturellen Modelle können als molekulare Grundlage für weitere Versuche, die durch *Ctr1* vermittelte Aufnahme von Cisplatin zu verstehen, dienen.

(ii) Dann haben wir vorläufige Berechnungen zur Wechselwirkung zwischen einem Cisplatin-DNA-Addukt und dem High-Mobility-Group-Box-1-(*HMGB1*)-Protein durchgeführt. Es wurde gezeigt, dass das *HMGB1* (ein DNA-bindendes Protein) in der Lage ist, den NER Reparaturweg durch Bindung an die Cisplatin-DNA-Addukte und durch deren Abschirmung, welche die Reparatur mit NER Proteinen verhindert, zu blockieren. Obwohl experimentelle Studien die strukturellen Bestimmungsparameter der Cisplatin-DNA-*HMGB1*-Komplexe aufgezeigt und auch Bindungsaffinitäten geliefert haben, ist der Mechanismus der molekularen Erkennung zwischen dem Cisplatin-DNA-Addukt und *HMGB1* noch unbekannt. Wir zielen darauf ab, strukturelle und energetische Erkenntnisse hinsichtlich der Bindung von *HMGB1* an das Cisplatin-DNA-Addukt zu gewinnen,

indem auf der Metadynamik Methode beruhende freie Energie Berechnungen durchgeführt werden. Diese Arbeit präsentiert die für dieses Ziel notwendigen Berechnungen.

Acknowledgments

First of all, I would like to deeply thank my supervisor Prof. Dr. Paolo Carloni for his support and guidance, which are essential to the completion of the thesis.

I would like to express my gratitude to Prof. Giovanni Natile and Dr. Fabio Arnesano at Bari University, Italy for their great collaboration. My PhD project has been benefited greatly from the important scientific insights of Prof. Natile and Dr. Arnessano. I would like to acknowledge that the EXAFS measurements shown in chapter 4 were done by Prof. Giovanni Natile's group.

Special thanks also go to Dr. Emiliano Ippoliti, Dr. Giulia Rossetti and Dr. Jens Dreyer for their patient help and suggestions throughout my PhD years in the Computational Biophysics Group.

I would like to extend my sincere thanks to all members of the Computational Biophysics Group for their help, discussions and for making me feel being at home at work.

I would like to thank the PhD committee: Prof. Dr. Evert Jan Meijer, Prof. Dr. Jörg Fitter and Prof. Dr. Carsten Honerkamp for their valuable time in reviewing my thesis work.

The financial support from German Research School for Simulation Sciences GmbH is sincerely acknowledged.

Finally, I would like to thank my family: Ba, Me, Anh Chi Hai, Hot Mit, Hot Dua and Nhan for their love, encouragement, patience, and trust in me.

Contents

1 Introduction **1**

2 Biophysical background **5**

2.1 Introduction 5

2.2 Mechanism of action 6

2.2.1 Cisplatin-DNA adducts 7

2.2.2 Recognition of cisplatin-DNA adducts: HMGB1 protein 8

2.3 Mechanisms of Resistance 10

2.4 Copper transport proteins and their binding to cisplatin 12

2.4.1 Ctr1 12

2.4.2 ATP7A, ATP7B and Atox1 13

3 Methods **17**

3.1 Molecular dynamics 17

3.1.1 Statistical mechanical basis of molecular dynamics 17

3.1.2 Derivation of molecular dynamics equation 18

3.1.3 Ab initio molecular dynamics 21

3.1.3.1 Ehrenfest MD 21

3.1.3.2 Born-Oppenheimer MD 22

3.1.3.3 Car-Parrinello MD 23

3.1.3.4 Density Functional Theory 24

3.1.3.5 Basis set 27

3.1.4 Force field-based molecular dynamics 29

3.1.5 Hybrid Car-Parrinello/molecular mechanics (QM/MM) molecular dynamics 30

3.1.6 Force matching procedure 32

3.1.7 Ewald summation . . . 33
3.1.8 Molecular dynamics in NPT ensemble . . . 34
3.1.8.1 Thermostats . . . 35
3.1.8.2 Barostats . . . 37
3.1.9 Replica exchange with solute tempering . . . 39
3.2 Computational spectroscopy . . . 40
3.2.1 NMR chemical shifts . . . 41
3.2.2 Extended X-ray absorption fine structure (EXAFS) . . . 43
3.2.3 Circular dichroism (CD) . . . 45

4 Structural determinants of cisplatin and transplatin binding to Ctr1's Met-rich motif: a computational spectroscopy approach 49

4.1 Introduction . . . 50
4.2 Computational protocol . . . 51
4.2.1 Molecular models . . . 51
4.2.2 QM/MM simulations . . . 53
4.2.3 Force matching procedure . . . 54
4.2.4 Replica exchange with solute tempering (REST) simulations . 54
4.2.5 Computational spectroscopy . . . 55
4.2.5.1 CD . . . 55
4.2.5.2 NMR . . . 56
4.2.5.3 EXAFS . . . 57
4.3 EXAFS measurements . . . 57
4.4 Results and discussion . . . 58
4.4.1 Peptide conformation . . . 58
4.4.2 Platinum coordination . . . 59
4.4.2.1 EXAFS spectra . . . 62
4.4.2.2 NMR chemical shifts . . . 63
4.5 Conclusions . . . 65

5 Binding of the platinated DNA to HMGB1 protein: preliminary results **69**

5.1 Introduction 69

5.2 Derivation of force field parameters for the platinated moiety 71

5.2.1 QM/MM simulation 71

5.2.2 Force matching procedure 72

5.3 Force field based MD simulation 73

5.4 Summary 74

6 Summary and outlooks **81**

A Appendix **83**

A.1 DFT-based geometry optimization of the model compounds 83

A.2 Conversion of NMR chemical shifts 83

A.3 QM/MM simulations and ^{195}Pt NMR chemical shielding calculation of the reference compound $[PtCl_6]^{2-}$ 84

A.4 Sampling of the QM region by QM/MM simulations 84

A.5 Metadynamics method 85

A.6 Supporting tables and figures 87

Bibliography **95**

List of Figures **116**

List of Tables **120**

Curriculum Vitae **123**

1

Introduction

The discovery of the pharmacological properties of cisplatin (Fig. 1.1a) has been an utmost important event in cancer research [1, 2, 3, 4, 5, 6]. The approval of cisplatin by U.S. Food and Drug Administration back in 1978 revolutionized anticancer therapy [7, 8]. Today, cisplatin is one of the most important anticancer drugs [1, 2]. It is widely used in clinics to treat various types of cancers, such as testicular, ovarian, cervical, head and neck, esophageal, and small-cell lung cancers [1, 2, 4, 8]. Cisplatin is especially effective in the treatment of testicular cancer with the overall cure rate exceeding 90% [4, 7]. In contrast with cisplatin, its geometrical isomer, transplatin (Fig. 1.1b), is inactive against cancers.

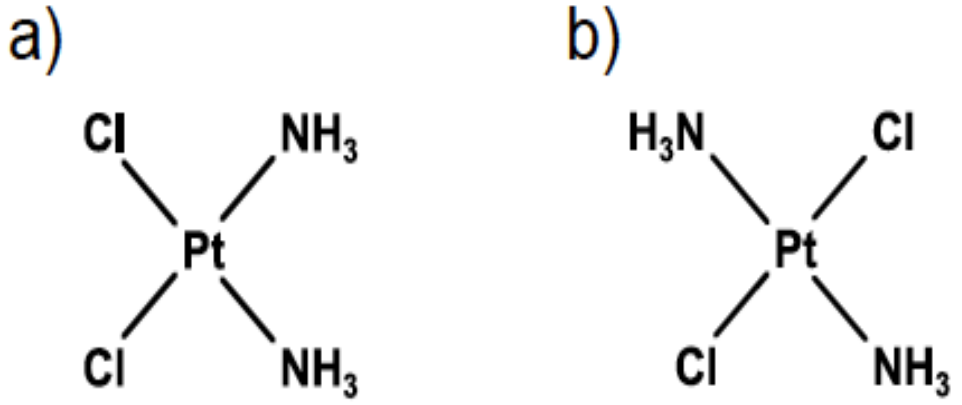


Figure 1.1 Chemical structures of a) cisplatin (*cis*-diamminedichloroplatinum(II)) and b) transplatin (*trans*-diamminedichloridoplatinum(II)).

DNA is believed to be the primary target of cisplatin [1, 2, 3, 4, 9]. Cisplatin kills cancer cells by covalently binding to their DNA at adjacent guanine (G) bases and causing the formation of intra-strand cross-links in the DNA (called cisplatin-DNA adducts) [9, 12, 20]. The cisplatin-DNA adducts in turn inhibit DNA replication [21, 22] and transcription [24, 25]. Since these are extremely important cellular

processes, their inhibition causes tumor cells to undergo suicide via the so-called apoptosis process (programmed cell death) [2, 28].

To reach its target, cisplatin takes a long journey, from the blood stream to the tumor cell's nucleus, where the DNA is located. On that route, cisplatin binds to a variety of proteins (Fig. 1.2). This causes several major drawbacks. First, protein binding may lead to harmful side effects [9, 29, 143]. Hence, cisplatin is toxic. Second, binding to proteins is exploited by the tumor cells for the development of drug resistance [2, 4, 28, 30, 85, 86]. The latter is correlated to a decrease in the rate of drug influx and to an increase in the rate of drug efflux/deactivation, with respect to non-resistant cells [2, 30, 82]. It follows that only a small fraction of the drug given to patients (about 1%) reaches DNA [30, 87]. Furthermore, drug-resistant cells are able to repair cisplatin-induced DNA damages [2, 4, 28]. They use a number of proteins involved in the so-called Nucleotide Excision Repair (NER) pathway to remove platinated nucleotides from the DNA, allowing cancer cells to survive [2, 28, 86, 98]. Second generation platinum-based drugs currently used in clinics such as oxaliplatin and carboplatin essentially suffers from the same problems [2, 4, 14, 24]. Understanding and eventually counteracting drug toxicity and resistance require a detailed knowledge of the molecular events associated with cisplatin's cellular uptake and platinated DNA repair mechanisms [28, 29, 86, 110]. What we have known so far originates exclusively from experimental studies and it is summarized here [29, 30, 31]:

(i) Cisplatin enters the cell by passive diffusion [30, 31] and by trans-membrane proteins, such as the copper transporter 1 (*Ctr1*) [29, 30, 110] (Fig. 1.2).

(ii) Once inside the cell, cisplatin can follow rather different fates (Fig. 1.2):

Deactivation: Cisplatin can be deactivated by binding to many proteins, including the thiol-rich metallothioneins (MTs), glutathione (GSH) [32] and the copper chaperone *Atox1* [33, 34].

Increase of efflux: The copper efflux proteins *ATP7A* and *ATP7B* in the secretory vesicle have been found to bind cisplatin and enhance its efflux [33].

Binding to DNA: Only a small fraction of cisplatin, which evades the inactivation and efflux processes above, enters the nucleus by an as-yet undetermined mechanism and interacts with DNA [30]. There, NER repair mechanism may take place. Some proteins, such as the High Mobility Group Box 1 (*HMGB1*) protein [1, 24, 76, 78], have been found to inhibit the NER repair mechanism by binding specifically to the cisplatin-DNA adduct and shielding it from being repaired by NER proteins [72, 73, 74].

Structural information on the cisplatin-protein adducts involved in all of these processes is essential to improve our understanding of drug resistance mechanism and to design new drugs with reduced resistance and increased ability to bind to DNA. Unfortunately, such information is mostly lacking, with a notable exception of the complex between *HMGB1* and platinated DNA [76].

Computational approaches are widely used to investigate cisplatin and other platinum drugs and their interactions with biomolecules [45]–[62], [155]–[165]. However,

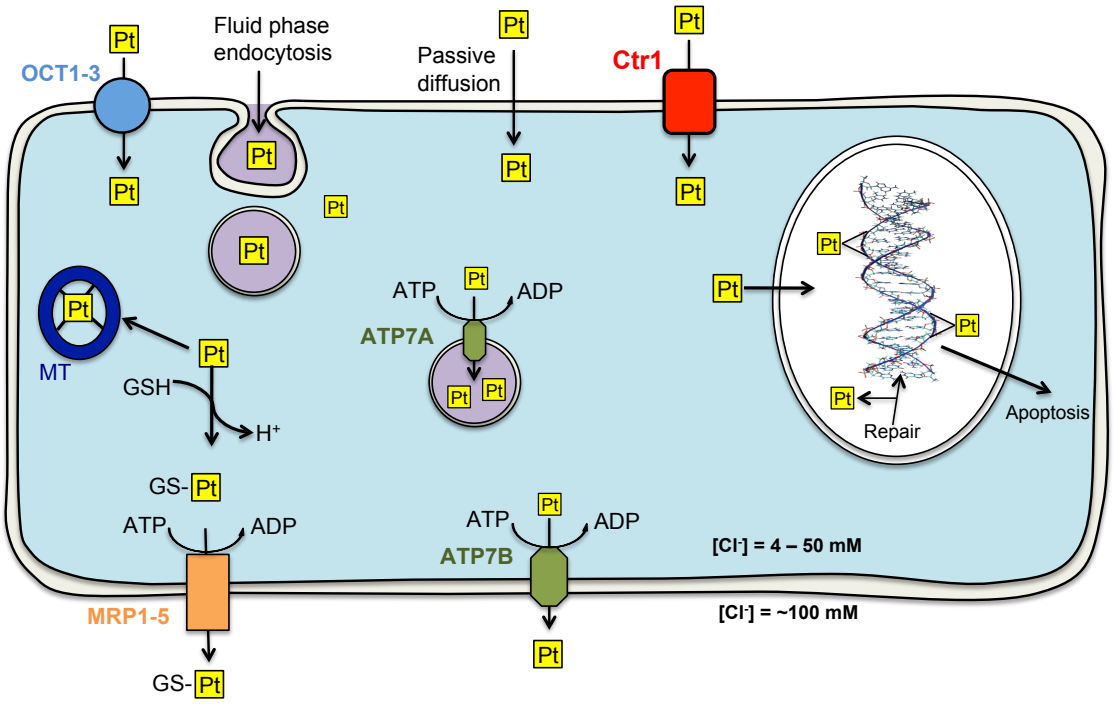


Figure 1.2 A simplified scheme of the mechanisms affecting cellular uptake, distribution and efflux of cisplatin (adapted from [30]).

almost all of the computational studies so far have focused mainly on structural properties of adducts of cisplatin and other platinum drugs with DNA [45]–[62]. The key issues of toxicity and resistance of cisplatin have been much less investigated computationally. This thesis is meant to provide one of the first contributions in this field. It focuses on two aspects:

(i) Binding of cisplatin and its isomer transplatin to a copper transport protein called *Ctr1*. *Ctr1* is a trans-membrane protein, located on the cell's surface [111, 112, 113, 114, 115] (Fig. 1.2). It has been found to transport cisplatin across the cell's membrane [29, 30, 110] although the exact transport mechanism is still unknown. Here, using computational approaches, we have provided the molecular structures of complexes of cisplatin and transplatin with a peptide (called *Mets7*) that mimics the extracellular binding domain of *Ctr1* [63]. Validation of our numerical results by biophysical experiments is fundamental. Therefore, the computational protocols presented in this thesis have been chosen to make comparison with available experimental data as direct as possible. Starting from molecular simulations we have computed NMR, EXAFS and CD spectra, which have been systematically compared with the corresponding experimental data. This approach represents a significant improvement with respect to the computational studies published so far, dealing essentially with the comparison between computed and experimental structures [45]–[62], [155]–[165]. Indeed, the computation of spectroscopic data allows a more direct, hence, reliable comparison with experiment. The molecular structures arising from our simulations can serve as the molecular basis for further attempts to elucidate the mechanism of cellular uptake of cisplatin.

(ii) Investigation of the mechanism of molecular recognition between cisplatin-DNA adduct and *HMGB1* protein. The free energy profile as a function of reaction coordinates describing the association process of *HMGB1* protein and cisplatin-DNA adduct is a key physical quantity to elucidate the molecular mechanism of the binding process. The free energy difference can be quantitatively correlated with the experimental binding affinity [283]. This thesis presents preliminary calculations, which will serve as a basis to perform free energy calculations using metadynamics method [284, 285].

2

Biophysical background

2.1 Introduction

The antitumor effect of cisplatin was discovered by chance by Barnett Rosenberg at Michigan State University in 1965 [5, 6, 8]. The drug was approved for clinical use by the U.S. Food and Drug Administration (ADF) in 1978 [8]. Cisplatin is an inorganic compound, in which a Pt(II) ion is coordinated by two ammine and two chlorido ligands in *cis* square planar geometry (Fig. 1.1a). Its geometrical isomer, transplatin has the same chemical composition, but the ligands are in *trans* arrangement (Fig. 1.1b). While cisplatin is active against many tumors, transplatin is not [10].

Nowadays, cisplatin is routinely used in clinics either as a single agent or in combination therapy to treat various types of cancers such as testicular, ovarian, small-cell lung, head and neck, cervical, and colorectal cancers [1, 2, 11]. Cisplatin is particularly effective against testicular cancer, for which the overall cure rate exceeds 90%, or nearly 100% for early stage disease [7].

However, cisplatin suffers from two major shortcomings: toxic side effects and resistance [1, 4, 28, 13]. Treatment with cisplatin indeed causes severe side effects such as nephrotoxicity, nausea and vomiting, bone marrow damage and neurotoxicity [13]. Moreover, many tumor cells displays intrinsic or acquired resistance to cisplatin, which further limits its clinical utility [2, 4]. These problems of cisplatin have generated a large interest in developing new platinum-based drugs with a better clinical profile [2, 3, 14]. This has led to the discovery and approval of second generation of platinum drugs, including carboplatin and oxaliplatin. However, these drugs also suffer from resistance [1, 2, 14]. The efforts to overcome cisplatin resistance have been proven very challenging and require a detailed knowledge of how cisplatin induces its antitumor effect and, more importantly, how tumors are or become resistant.

2.2 Mechanism of action

DNA binding is believed to be a key cellular event that triggers the anticancer activity of cisplatin [1, 2, 4, 9]. Studies over many years have provided convincing evidences in favor of this hypothesis [1]. Bacteria treated with cisplatin show phenotypes that are characteristic of those evoked by DNA-damaging agents [15]. In addition, DNA-repair-deficient cells are indeed more sensitive to cisplatin [16, 17]. There is a significant correlation between cisplatin-DNA adduct levels and sensitivity of cisplatin-treated cells [18, 19].

Cisplatin binds to DNA, preferentially at the N7 position of guanine bases [9, 20, 23], inhibiting replication [9, 21, 22] and transcription of the DNA [24, 25, 26, 27]. The inhibition of these critical DNA-related processes triggers subsequent intracellular events leading to apoptotic cell death [28]. However, before binding to the DNA, cisplatin has to survive a long journey from the blood stream to the tumor cell's nucleus where DNA is located. The current knowledge on the steps leading cisplatin from intravenous injection to nuclear DNA binding can be summarized as follows (Fig. 1.2) [29, 30, 31].

(i) Cisplatin diffuses in the blood stream and in the extracellular fluids, where the chloride (Cl^-) concentration is relatively high ($\sim$100mM) [29, 30]. The high concentration of chloride prevents cisplatin from undergoing aquation reactions, which transform cisplatin into activated monoaquated or diaquated species (*cis*-$[Pt(NH_3)_2H_2OCl]^+$ and *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$). Therefore, cisplatin can reach the cancer cells' surface in the intact neutral form [31].

(ii) Cisplatin enters cells both by passive diffusion across the cell membrane and by active transport using the high-affinity copper transporter 1 *Ctr1* and the organic cation transporters *OCT*1–3 [30]. Pinocytosis, characterized by the formation of an endocytic vesicle incorporating a portion of external solution containing all chemicals there present (including cisplatin), is another possibility of drug uptake [132] (Fig. 1.2).

(iii) Once inside the cell, the lower concentration of chloride in the cytoplasm (4–50mM) facilitates the aquation process of cisplatin to form *cis*-$[Pt(NH_3)_2H_2OCl]^+$ and *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ [30, 31]. The aquated forms are more reactive towards cellular targets. Although DNA is the primary target of cisplatin, many other nucleophilic components present in the cytosol can bind cisplatin. These include thiol-rich metallothioneins (MTs), glutathione (GSH) [32] and the copper chaperone *Atox1* [33, 34]. All of these proteins sequester cisplatin, and eventually deactivate it. Cisplatin can also be exported out of the cell by the copper pumps *ATP7A* and *ATP7B* in the secretory vesicles. These proteins can bind and enhance cisplatin efflux [33]. Finally, only a small fraction of cisplatin, which evades the inactivation and efflux processes above, enters the nucleus by an as-yet undetermined mechanism and bind to DNA, triggering apoptotic cell death if the DNA lesion is not repaired [30, 31].

2.2.1 Cisplatin-DNA adducts

Cisplatin forms a spectrum of intra- and inter-strand DNA cross-links, which have been identified both *in vitro* and *in vivo* (Fig. 2.1) [1, 24, 35, 36, 37, 38]. The major adduct, comprising ∼65% of the total products, is a *cis*-1,2-$[Pt(NH_3)_2]^{2+}$-d(GpG) intra-strand cross-link (Fig. 2.1a) [1]. Other minor products include 1,2-d(ApG) (∼25%) and 1,3-d(GpNpG) (Fig. 2.1b) (5–10%) intra-strand adducts, as well as a smaller number of inter-strand cross-links (Fig. 2.1c) and monodentate adducts [1]. Transplatin, a clinically inactive isomer of cisplatin (Fig. 1.1b) is unable to form 1,2-intra-strand cross-links, due to stereo-chemical constraints [39, 40]. This implies that intra-strand adducts are more likely to be responsible for the anticancer activity of cisplatin [1]. There is no correlation between the frequency of inter-strand cross-links formation and cytotoxicity of cisplatin, which provides further evidence that intra-strand cross-links are essential to tumor cell death [1, 39, 40].

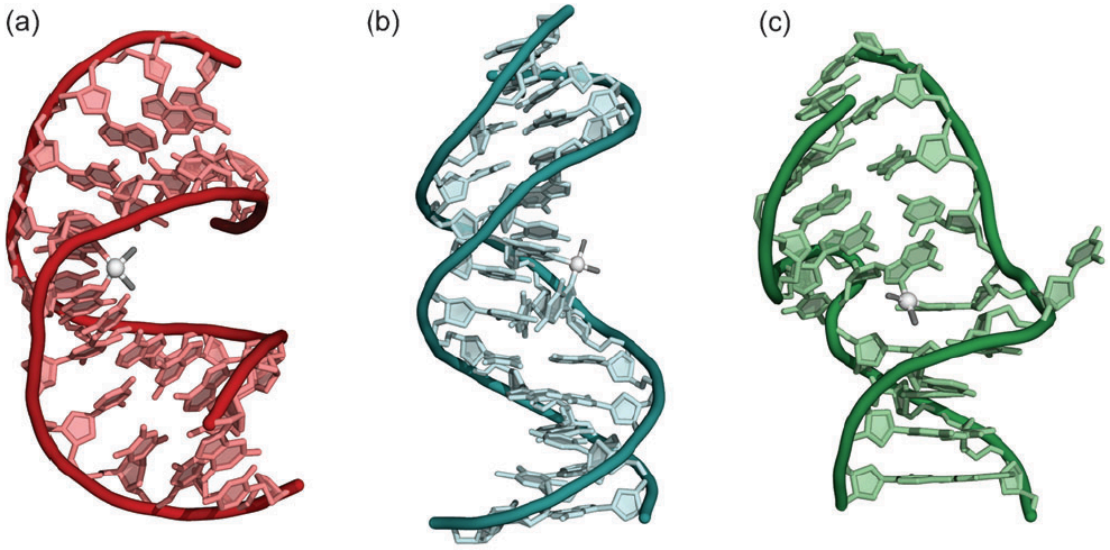


Figure 2.1 Major cisplatin-DNA adducts: (a) 1,2-$[Pt(NH_3)_2]^{2+}$-d(GpG) intra-strand cross-link; (b) 1,3-$[Pt(NH_3)_2]^{2+}$-d(GpTpG) intra-strand cross-link; (c) inter-strand cross-link (adapted from [24]). DNA's backbone is shown in ribbon; their nucleobases in bond representation; the platinum atom is represented as a sphere.

The X-ray structure of a platinated DNA dodecamer duplex containing the major *cis*-1,2-$[Pt(NH_3)_2]^{2+}$-d(GpG) intra-strand cross-link revealed that DNA platination induces a global bend in the DNA duplex by 35°–40° and unwinds the double helix by ∼25° [1, 24, 41, 42] (Fig. 2.2a). The major groove is compacted and the minor groove widened and flattened. The DNA takes on A-form properties to the 5' and a B-form structure on the 3' side of the 1,2-d(GpG) adduct. The roll angle between platinated guanine bases is ∼26° [41]. NMR spectroscopic studies [1, 43, 44] revealed some differences between the solid state and solution structures. The solution structures (Fig. 2.2b) shows a bend angle of 60°–70° and a large roll angle of ∼49° at the 1,2-$[Pt(NH_3)_2]^{2+}$-d(GpG) intra-strand cross-link. In addition, the NMR structures contain exclusively B-form DNA. The 1,3-$[Pt(NH_3)_2]^{2+}$-d(GpTpG)

intra-strand cross-link on the duplex DNA (Fig. 2.1(b)) has also been solved by NMR spectroscopy [64, 65]. In this structure, the duplex is bent by ∼30° and the double helix displays local unwinding and widening of the minor groove, similarly to features of the 1,2-$[Pt(NH_3)_2]^{2+}$-d(GpG) intra-strand cross-link. However, the 1,3-$[Pt(NH_3)_2]^{2+}$-d(GpTpG) adduct differs in that base pairing of the 5' G*-C, where the "*" denotes the platinated site, is disrupted and the internal thymidine of the adduct is extruded from the minor groove [24].

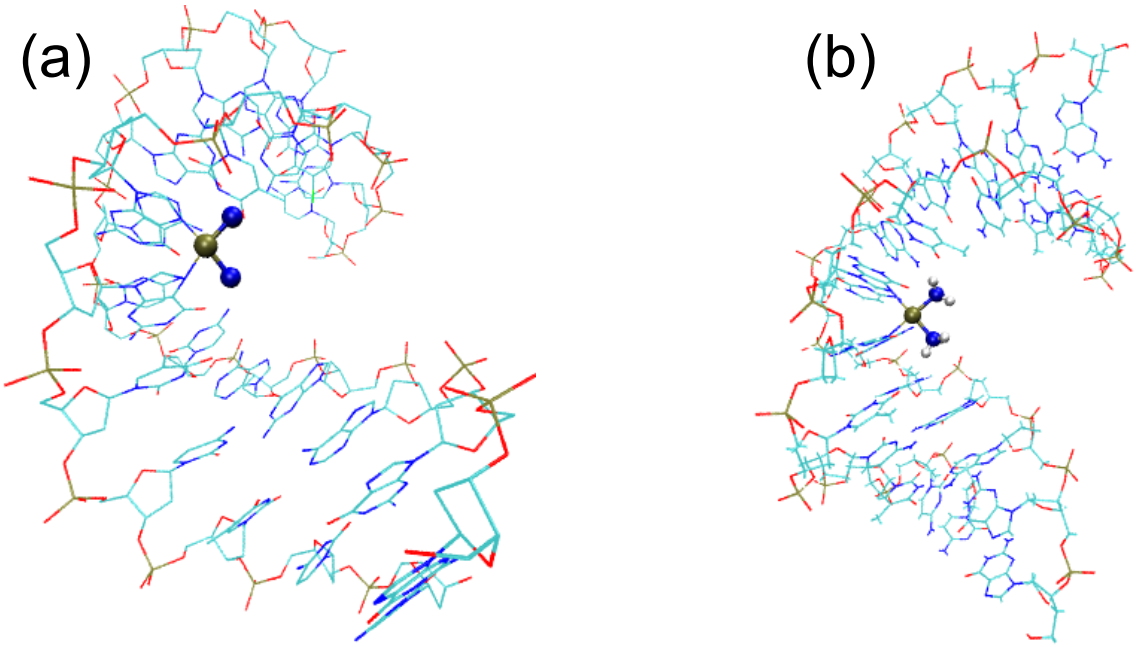


Figure 2.2 Structures of cisplatin-DNA adduct: (a) from X-ray [41] and (b) from NMR [44]. Atoms in DNA are represented by lines while the cisplatin moiety is shown in ball-stick model (platinum in gray, nitrogen in blue).

2.2.2 Recognition of cisplatin-DNA adducts: HMGB1 protein

Many proteins have been identified to bind to cisplatin-DNA adducts [1, 9, 24]. These include High Mobility Group Box (*HMGB*) proteins, repair proteins, transcription factor proteins and others [1, 3, 24]. Among these proteins, the High Mobility Group Box 1 (*HMGB1*) protein has been extensively studied, and for which most information is available.

HMGB1 protein is one of the early proteins discovered to bind cisplatin-DNA adducts [66, 67]. *HMGB1* is an abundant protein (∼106 copies per cell) [68]. It has been suggested that *HMGB1* play an important role in regulation of many cellular processes involving DNA, such as, chromatin remodeling, recombination, replication, and transcription [68, 69]. *HMGB1* expression levels have been shown to correlate with cisplatin sensitivity. *In vitro* repair essays revealed that *HMGB1* can inhibit Nucleotide Excision Repair (NER) of cisplatin-DNA adducts (see Sec. 2.3), presumably by binding to and shielding the damaged site from recognition by the repair

proteins [70, 71, 72, 73, 74]. An increased level of *HMGB1* following hormone treatment sensitizes breast cancer cells to cisplatin by 2-fold (the number of cancer cells killed by cisplatin increases twice when treated with the same amount of drug) [74]. Moreover, additional expression of *HMGB2*, a protein over 85% identical to *HMGB1*, in human lung cancer cells enhanced cisplatin sensitivity more than 3-fold [75].

The structure of a complex between 16mer duplex DNA containing a 1,2-$[Pt(NH_3)_2]^{2+}$-d(GpG) intra-strand cross-link and the domain A of *HMGB1* was solved by X-ray crystallography (Fig 2.3) [76]. *HMGB1* domain A contains three alpha helices forming an L-shape structure that facilitates DNA binding [77]. The concave surfaces of *HMGB1* domain A form hydrogen bonds as well as hydrophobic and water-bridging contacts with the minor groove of the platinated DNA. Phe37 intercalates into the hydrophobic cleft formed by the two platinated guanines bases G8 and G9 (Fig. 2.3) [76]. The binding affinity between *HMGB1* and cisplatin-DNA adduct was also measured ($K_d = 0.7 \pm 0.1$ nM) (see Sec. 5.1 for the definition of K_d) [78].

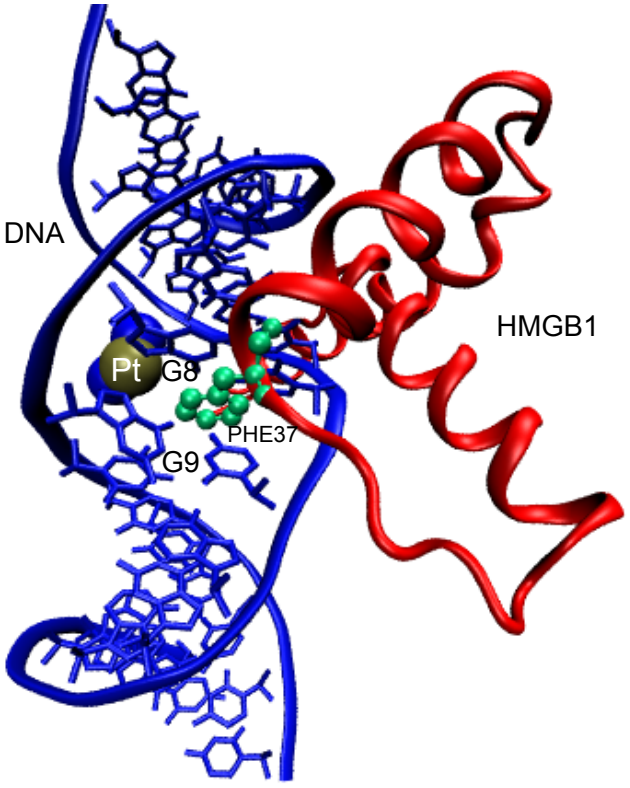


Figure 2.3 X-ray structure of the complex between *HMGB1* domain A and *cis*-[Pt(II)$(NH_3)_2$]-d-(CCUCTCTG*G*ACCTTCC)-d(GGAGAGACCTGGAAGG) (PDB code 1CKT) [76]. The DNA's backbone is shown in blue ribbon; nucleobases in blue bonds; cisplatin moiety in van der Waals balls. The *HMGB1*'s backbone is shown in red ribbon; its Phe37 amino acid, which inserts between two platinated G8 and G9 bases, is shown in green ball-stick representation.

2.3 Mechanisms of Resistance

The development of resistance is a common drawback of chemotherapy, and cisplatin is no exception [4, 79]. It is a major cause of treatment failure [4]. Resistance implies that tumor cells fail to undergo apoptosis (programmed cell death) at a clinically relevant drug concentration [28]. Some tumors, such as colorectal cancer and non-small-cell lung cancer are intrinsically resistant to cisplatin chemotherapy [80]. Other tumor types, such as head and neck cancer, testicular cancer, ovarian cancer and small-cell lung cancer are predominantly sensitive to cisplatin treatment. However, most of these sensitive tumors develop acquired resistance after the initial treatment [4]. The mechanisms by which tumor cells develop resistance to cisplatin are multifactorial, and not completely understood [3, 28, 30, 81, 85]. The generally accepted mechanisms [2, 28, 30, 85, 86] by which cancer cells acquire resistance to cisplatin are: (i) decreased accumulation of cisplatin, (ii) increased deactivation of cisplatin by thiol-containing molecules and (iii) increased repair of DNA adducts.

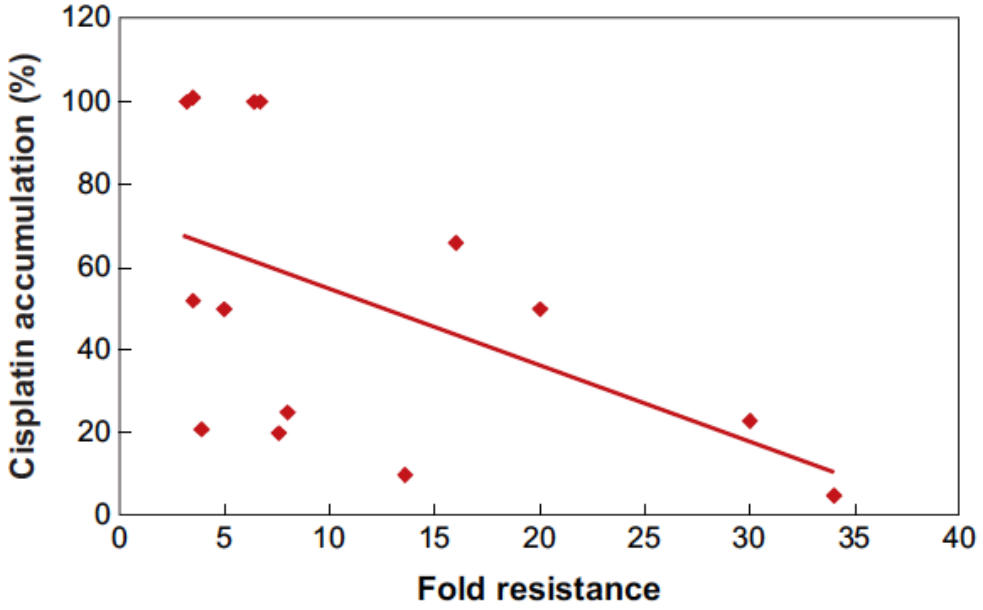


Figure 2.4 Correlation between the fold-resistance (which is defined as the fold increase in the number of survival cells when treated with the same amount of drug) and the percent decrease in cisplatin accumulation with respect to sensitive cells (adapted from [30]).

(i) **Decreased accumulation of cisplatin:** This is the most common and distinctive feature of resistant cells [28, 30]. It has been observed repeatedly in different types of tumor cells that cisplatin accumulation in resistant cells is reduced significantly in comparison to that in sensitive cells [2, 30, 82]. Resistant cells have apparently altered the cellular transport machinery so as to reduce the amount of cisplatin to reach the target DNA. Over many years, several studies have been devoted to investigate the ability of cisplatin and other platinum drugs to accumulate in cancer cells. It has been consistently

demonstrated that accumulation of the drug is a determinant of cellular sensitivity to cisplatin [83, 84]. In most of human- and rodent- cell lines, which are resistant to cisplatin, its accumulation is reduced relative to the non-resistant cells [30, 87]. A strong correlation between cisplatin accumulation and relative cisplatin resistance were observed for a series of increasing resistant cell lines derived from the BEL7404 human hepatoma cell line (Fig. 2.4) [88]. Moreover, Koga *et al.* examined seven primary bladder cancer cell lines derived from untreated cancer of urinary bladder and found a positive correlation between cisplatin accumulation and sensitivity among the intrinsically resistant cell lines [30, 89]. Diminished accumulation has been observed not only for cisplatin but also for carboplatin [90] and oxaliplatin [91]. Reduced intracellular accumulation of cisplatin is due to either decreased uptake or increased efflux or both. It is, therefore, important to understand the molecular mechanisms of cellular uptake and efflux of cisplatin.

(ii) **Increased deactivation of cisplatin by thiol-containing molecules:** Thiol (SH)-containing molecules such as intracellular tripeptide glutathione (GSH) or metallothioneins (a class of low-molecular weight proteins) are able to bind and deactivate cisplatin [4, 28, 92]. GSH reacts with cisplatin and other electrophilic compounds to form deactivated conjugates that are readily excreted by a GS-conjugated export pump [93]. Early studies using eight human ovarian carcinoma cell lines showed a significant correlation between sensitivity to cisplatin, and levels of GSH [28, 94]. A study of two ovarian cancer cell lines that were derived from the same patient before and after the onset of drug resistance showed 2.9-fold higher levels of GSH in the cells derived from the drug-resistant tumor [95]. Over-expression of metallothioneins, that are involved in heavy-metal binding and detoxification, in either mouse cells [96] or in human ovarian cancer cells [97] led to 4-fold or 7-fold increase in cisplatin resistance, respectively.

(iii) **Increased repair of DNA adducts:** An important biochemical mechanism of resistance is the repair of cisplatin-DNA adducts [2, 4, 28, 86, 98]. Many cisplatin-resistant cell lines that are derived from various tumor types have been shown to have increased DNA-repair capacity in comparison to sensitive counterparts [99]. Nucleotide Excision Repair (NER) is a major mechanism of cisplatin-damaged DNA repair. Increased NER in cisplatin-resistant cell lines has been shown to occur both for intra-strand as well as inter-strand cisplatin-DNA adducts [100, 101]. NER is an ATP-dependent multi-protein complex that recognizes the kink induced on the DNA by 1,2-intrastrand cross-link and subsequently excises the segment containing the kink. The gap is then filled by DNA polymerase [102]. Increased removal of DNA adducts by NER in resistant cell lines relative to sensitive parent cell lines has been consistently observed in several models [9, 103]. Conversely, defective NER has been found in cell lines with hypersensitivity to cisplatin [104].

2.4 Copper transport proteins and their binding to cisplatin

Reduced cellular accumulation of cisplatin is a major mechanism of drug resistance [2, 30, 82]. Therefore, it is vitally important to understand how the cellular transport machinery of cisplatin works. For a long time, cisplatin was believed to enter cells only by passive diffusion [105, 106, 107, 108, 109]. Although passive diffusion mechanism has been not fully ruled out, recently there has been a growing body of evidence indicating that cisplatin as well as other platinum drugs enter cells, are distributed to various sub-cellular compartments, and exported from the cells via transport proteins that evolved to manage copper homeostasis [29, 30, 110, 131, 132]. In particular, the high-affinity copper transporter 1 (*Ctr1*) has been shown to be involved in the uptake of cisplatin and other platinum drugs across the cells' membrane, whereas the ATPase copper pumps *ATP7A* and *ATP7B* are responsible for sequestration and efflux of cisplatin [30].

2.4.1 Ctr1

Ctr1 is an evolutionarily conserved copper uptake protein present in plants, yeast and mammals. It is the main copper importer in mammalian cells [111, 112, 113, 114, 115]. *Ctr1* in human is expressed in all tissues. It plays a key role in homeostatic regulation of cellular copper levels. Recently, *Ctr1* has been demonstrated to regulate the cytotoxicity of cisplatin by affecting drug uptake [116, 117, 118, 119, 120, 121]. The absence of *Ctr1* expression causes yeast and mammalian cells resistant to cisplatin, carboplatin and oxaliplatin [116, 117, 118, 119, 122]. On the other hand, over-expression of *Ctr1* was shown to sensitize cells to the toxic effects of cisplatin [118, 123]. These changes in sensitivity are accompanied by proportionate changes in the uptake of cisplatin and other platinum agents [110]. *In vivo* studies in rats showed that cisplatin cause atrophy in those cells in dorsal root ganglia that expressed high levels of *Ctr1* but not those with low-level expression [124]. It was recently demonstrated that knockout of *Ctr1* completely eliminates responsiveness of tumors to cisplatin *in vivo* in a murine model [122].

Molecular aspects:

Ctr1 is a membrane protein having three trans-membrane domains. A homotrimer of *Ctr1* forms a pore in the lipid bilayer (Fig. 2.5). The extracellular domain of *Ctr1* caries several methionine-rich motifs, which have been identified as a putative cisplatin binding site [142].

To provide insights on interactions of *Ctr1* with cisplatin, transplatin and other platinum complexes, Arnesano *et al.* chose one of the Met-rich motifs of yeast *Ctr1* (called *Mets7*) with slight modification [144] to make it more water-soluble and performed the reaction with cisplatin, transplatin and other platinum complexes [142, 143]. The investigation was performed using mainly electrospray ionization mass spectroscopy (ESI-MS), NMR and circular dichroism (CD).

CD measurements showed that *Mets7* features a β-turn conformation after reacting with cisplatin, whilst it is a random coil in the Apo form [142]. Transplatin was

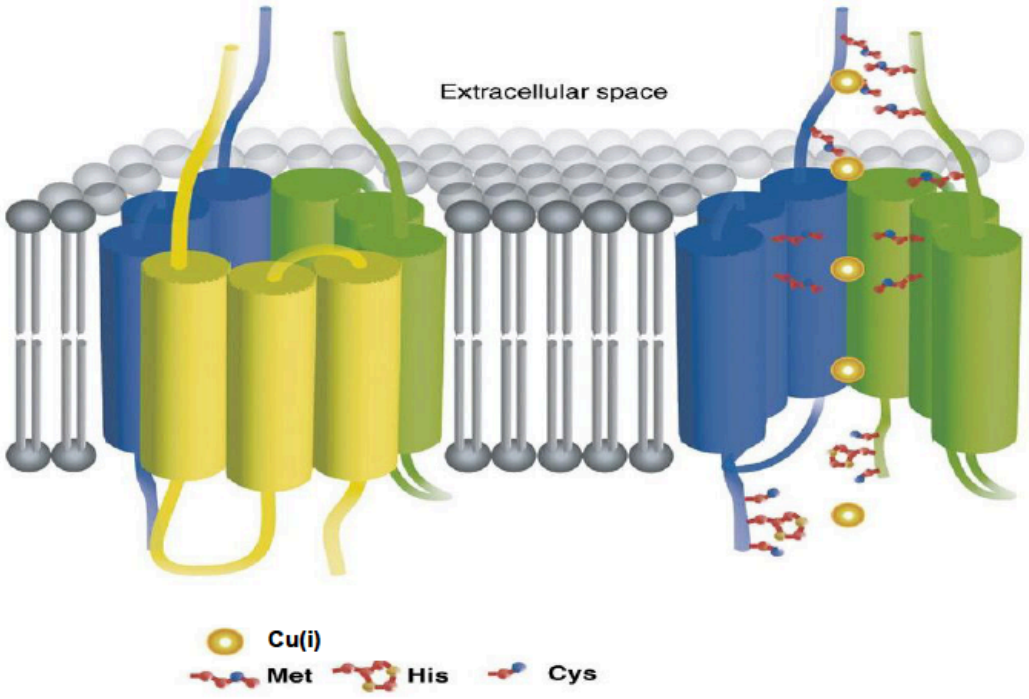


Figure 2.5 Cartoon representing the copper influx transporter *Ctr1* (adapted from [145]). Left: *Ctr1* is a symmetrical homotrimeric trans-membrane protein. Nine trans-membrane helices form a pore in the lipid bilayer. Each subunit is represented by a different color. Right: One of the three subunits is removed to reveal the interior of the pore.

shown to modify just slightly the random coil of *Mets7*. ESI-MS data showed that cisplatin progressively loses all ammine and chlorido ligands and ultimately coordinates only to *Mets7*. This contrasts with transplatin, which keeps the two ammine ligands [142, 143]. Consistently, $^{1}H/^{13}C$ heteronuclear single-quantum coherence (HSQC) spectra (Fig. 2.6) reveals a large downfield shift of the ^{1}H and ^{13}C nuclei of ϵ-CH_3 and γ-CH_2 of methionines and confirms coordination of Pt to the S atoms. Other groups also reported a similar binding mode of cisplatin to partial or full-length N-terminal domain of human *Ctr1* [146, 147]. Hence, *Mets7* is a good mimic of the whole N-terminal domain of *Ctr1* in terms of reactivity against cisplatin. The full structural determinants of cisplatin- and transplatin-*Mets7* adducts are so far lacking.

2.4.2 ATP7A, ATP7B and Atox1

ATP7A and *ATP7B* are functionally conserved P-type ATPases involved in the sequestration and extrusion of excess Cu ions [30, 113, 125, 126, 112]. The two proteins are homologous in structure, sharing 8 trans-membrane domains and ∼65% amino acid sequence identity. However, they are expressed in different tissues: *ATP7A* in intestinal epithelium and *ATP7B* in liver and kidney. Enhanced expression of *ATP7A* and *ATP7B* has been linked to the degree of cisplatin resistance in tumor cell lines and clinical samples [128]. Over-expression of *ATP7B* is associated with

a reduced capacity in cells to accumulate cisplatin and show a high rate of efflux [129]. A confocal microscopy study showed that an ECFP (enhanced cyan fluorescent protein)-tagged *ATP7B* (*ATP7B*-ECFP) colocalized with a fluorescence labeled cisplatin (F-cisplatin) and that both of cisplatin and F-cisplatin caused trafficking of the *ATP7B*-ECFP containing vesicles toward the cell's surface [130]. This is consistent with the role of *ATP7B* as an efflux transporter of cisplatin. On the other hand, the function of *ATP7A* appears to be different from that of *ATP7B* [132]. Although *ATP7A* over-expressing cells exhibit a significant degree of enhanced resistance to cisplatin, higher levels of cisplatin accumulation were observed [133]. This indicates that the function of *ATP7A* is likely to be in binding and sequestration of cisplatin rather than exporting it [132, 133]. This was confirmed in many observations including detection of significantly higher levels of cisplatin in the vesicular compartment of the cells that over-express *ATP7A* [133]. Confocal microscopy demonstrated that *ATP7A* and F-cisplatin colocalized in the vesicular structures in ovarian carcinoma cells 2008 [134]. Exposure to cisplatin did not cause trafficking of the *ATP7A* containing vesicles toward the cell surface [133].

Atox1 is a key metallochaperone that receives copper from the importer *Ctr1* and delivers it to the copper exporters *ATP7A* and *ATP7B* in the secretory compartment [135, 136, 137, 112]. It has 69 amino acids and is found in the cystosol and nucleus [135]. Loss of *Atox1* in the fibroblast cells of mice has been found to reduce the influx of cisplatin and subsequent accumulation in vesicular compartment and DNA [138]. The exact role of *Atox1* in cellular transport and distribution of cisplatin is still not fully understood. However, it has been suggested that *Atox1* regulates the influx of cisplatin by controlling the cisplatin-induced down regulation of *Ctr1* through ubiquitination [110, 138]. *Atox1* has been found to translocate to the nucleus in response to copper exposure [139, 140] and it has also recently been reported to bind cisplatin [141, 34]. This raises the question of whether it could be involved in the delivery of cisplatin to DNA.

Molecular aspects:

Solution and in-cell NMR spectroscopy study [34] showed that cisplatin binds to the CxxC metal binding motif of *Atox1* (Fig. 2.7a). Initially, a *cis*-$[Pt(NH_3)_2]$-*Atox1* adducts is formed. At longer reaction time the protein dimmerization accompanied by loss of the two ammine ligands is observed [34]. This interaction mode is different from that observed previously by Boal *et al.* using X-ray crystallography technique [141].

ATP7B features six soluble domains, each of which contains about 70 amino acids. Their structures have been solved by NMR spectroscopy [148, 149]. Interestingly, each domain has the similar fold as the *Atox1* protein [150, 151]. All the domains share the same CxxC metal binding motif (Fig. 2.7b). A likely model of the adducts of cisplatin with *ATP7B* involves the binding of the *cis*-$[Pt(NH_3)_2]^{2+}$ moiety to the CxxC motifs [152] as suggested by cysteine mutation experiments [153, 154].

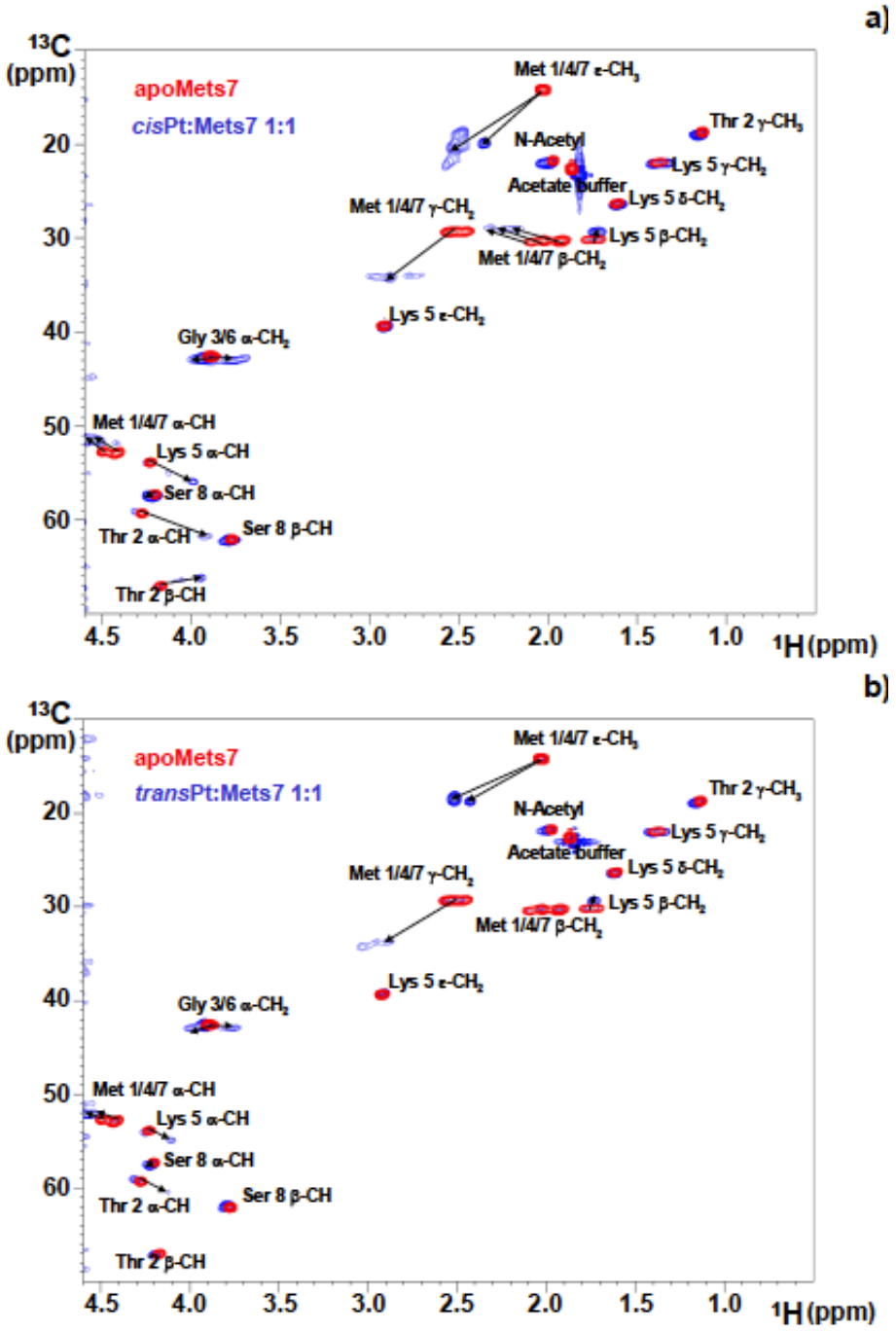


Figure 2.6 Overlay of 2D ^{1}H,^{13}C-edited HSQC spectra of *Mets7* (red contours) and of mixtures (1:1) of *Mest7* with a) cisplatin and b) transplatin. Spectral changes induced by cisplatin/transplatin are indicated with arrows. Figures adapted from [142].

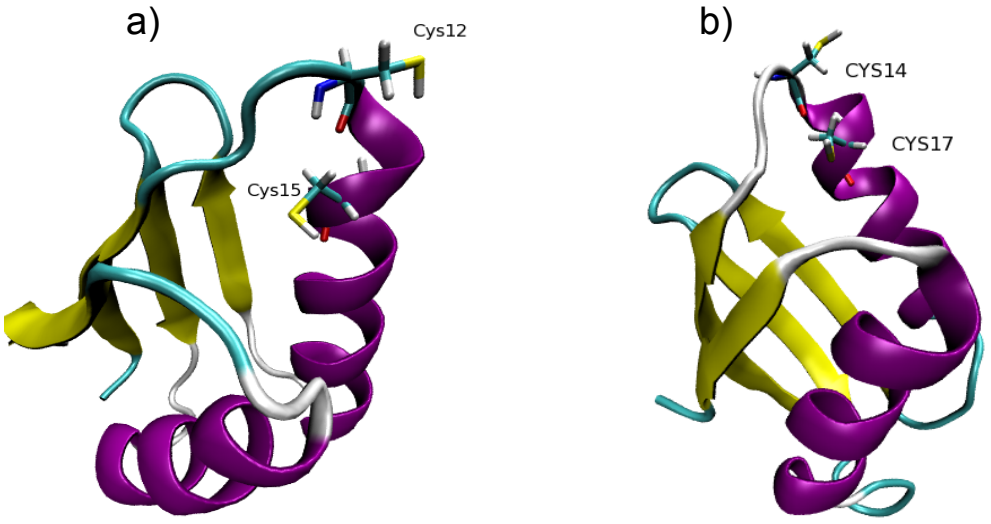


Figure 2.7 The NMR structures of: (a) *Atox1* [151] and (b): the soluble domain 4 of the *ATP7B* protein [148]. The backbone is represented with ribbon. Different colors indicate different secondary-structure types: green is loop, yellow is beta sheet, and purple is alpha helix. Two cysteine amino acids are shown in bond representation.

3

Methods

3.1 Molecular dynamics

3.1.1 Statistical mechanical basis of molecular dynamics

Molecular dynamics simulations generate information at the microscopic level (atomic positions, velocities...) and the connection of this very detailed information to macroscopic quantities (such as pressure, temperature, energy...) is provided by statistical mechanics [255]. The set of atoms' coordinates and momenta can be thought as coordinates in a multidimensional space: the phase space. For a system of N atoms, this space has $6N$ dimensions. Let us denote a particular point in the phase space as Γ, and suppose that we can write the instantaneous value of some property A as a function of Γ, $A(\Gamma)$. The experimentally observable value of A, A_{obs} is given by the ensemble average. Here, an ensemble is regarded as a collection of points Γ in the phase space. The points are distributed according to a probability distribution $p(\Gamma)$. The ensemble average (or thermodynamic average) is written as

$$A_{\text{obs}} = \langle A \rangle_{\text{ens}} = \int A(\Gamma) p(\Gamma) d\Gamma, \tag{3.1}$$

where $d\Gamma = d\vec{R}_1 \cdots d\vec{R}_N \cdot d\vec{P}_1 \cdots d\vec{P}_N$. The probability distribution depends on macroscopic parameters (NVE, NVT, NPT...), which defines the experimental conditions. In the case where N, V and T are kept fixed (canonical ensemble), the corresponding probability distribution is given by the Boltzmann distribution function

$$p_{NVT} = \frac{\exp\left[-H(\Gamma)/k_B T\right]}{Z}, \tag{3.2}$$

where

$$Z = \int d\Gamma p_{NVT}(\Gamma) = \int d\Gamma \exp\left[-H(\Gamma)/k_B T\right] \tag{3.3}$$

is the canonical partition function, k_B is the Boltzmann factor, $H(\Gamma)$ is the classical Hamiltonian of the system which is given by

$$H(\Gamma) = H\left(\left\{\vec{R}_I\right\}, \left\{\vec{P}_I\right\}\right) = \sum_{I=1}^{N} \frac{\vec{P}_I^2}{2M_I} + U\left(\left\{\vec{R}_I\right\}\right). \tag{3.4}$$

$\vec{R}_I$ and $\vec{P}_I$ are the coordinate and momentum of atom I, respectively. M_I is the mass of atom I. In order to calculate the ensemble average (Eq. (3.1)), it is necessary to calculate the partition function (Eq. (3.3)), which is an extremely difficult task.

Molecular dynamics (MD) offers a strategy to calculate the ensemble average. In this approach, the motion of a single point Γ through the phase space is followed as a function of time. The time evolution of a single point Γ through the phase space is obtained by integrating the equation of motion, starting from an initial point $\Gamma(0)$. The integration yields a trajectory that is a set of points $\Gamma(t)$ describing the state of the system at any successive time t. The "dynamical average" of A can be calculated by averaging over those points that have been visited

$$\langle A \rangle_\tau = \frac{1}{\tau} \int_0^\tau A\left[\Gamma(t)\right] dt. \tag{3.5}$$

The two averages (3.1) and (3.5) are connected by the "ergodic hypothesis". The system is at equilibrium and if it is allowed to evolve infinitely long in time, then it will pass through all its possible states. A system which obeys the conditions above is said to be ergodic, and under these conditions the dynamical average is equal to the ensemble average for an infinitely long trajectory

$$\lim_{\tau \to \infty} \langle A(\Gamma) \rangle_\tau = \langle A(\Gamma) \rangle_\Gamma . \tag{3.6}$$

The fact that all the "real" molecular systems are assumed to be ergodic is the theoretical justification for using MD simulations as a means to calculate ensemble averages of molecular systems [255]. In practice, the application of MD simulations to calculate thermodynamic averages assumes that the system under investigation is ergodic and that a finite simulation is a "good" approximation to the infinite trajectory.

3.1.2 Derivation of molecular dynamics equation

The time-dependent behavior of molecular systems is best described by the time-dependent Schrödinger equation. However, this equation is extremely difficult to solve for large systems. Therefore, classical mechanical description is often used to approximate the motion of atoms. In most of computational studies of biomolecules, the Newtonian equations of motion of the classical mechanics are solved, rather than the quantum mechanical equations. In this section, the classical equations of molecular dynamics (Newton's equations) will be derived from the time-dependent Schrödinger equation [257, 258, 259].

The time-dependent Schrödinger equation is written as

$$i\hbar \frac{\partial}{\partial t} \Psi\left(\{\vec{r}_i\}, \{\vec{R}_I\}; t\right) = H\Psi\left(\{\vec{r}_i\}, \{\vec{R}_I\}; t\right), \tag{3.7}$$

where H is the Hamiltonian of the system. For most of the biologically relevant systems, the non-relativistic Hamiltonian is sufficient to get a reasonably accurate description

$$\begin{aligned} H = & -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 \\ & + \frac{1}{4\pi\varepsilon_0}\sum_{i<j}\frac{e^2}{|\vec{r}_i - \vec{r}_j|} - \frac{1}{4\pi\varepsilon_0}\sum_{I,i}\frac{e^2 Z_I}{\left|\vec{R}_I - \vec{r}_i\right|} + \frac{1}{4\pi\varepsilon_0}\sum_{I<J}\frac{e^2 Z_I Z_J}{\left|\vec{R}_I - \vec{R}_J\right|} \\ = & -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{1}{4\pi\varepsilon_0}\sum_{i<j}\frac{e^2}{|\vec{r}_i - \vec{r}_j|} + V_{n-e}\left(\{\vec{r}_i\},\{\vec{R}_I\}\right) \\ = & -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + H_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right). \end{aligned} \tag{3.8}$$

$\{\vec{r}_i\}$ and $\{\vec{R}_I\}$ denote the coordinates of the electrons and nuclei, respectively. M_I and Z_I are the mass and atomic number of the I-th nucleus. m_e and $-e$ are the electron's mass and charge, respectively. ε_0 is the vacuum permittivity. $\Psi\left(\{\vec{r}_i\},\{\vec{R}_I\};t\right)$ is the total wavefunction.

In the one-determinant approximation ansatz [257, 258, 259], $\Psi\left(\{\vec{r}_i\},\{\vec{R}_I\};t\right)$ is separated into the electronic and nuclear contributions

$$\Psi\left(\{\vec{r}_i\},\{\vec{R}_I\};t\right) = \phi\left(\{\vec{r}_i\};t\right)\chi\left(\{\vec{R}_I\};t\right)\exp\left[\frac{i}{\hbar}\int_{t_0}^{t}\tilde{E}_e(t')dt'\right], \tag{3.9}$$

where the nuclear and electronic wavefunctions are separately normalized to unity at every instant of time, *i.e.*

$$\int \left|\phi\left(\{\vec{r}_i\};t\right)\right|^2 d\vec{r} = \langle\phi,t\,|\phi,t\rangle = 1, \quad \forall t\in\mathbb{R}, \tag{3.10}$$

$$\int \left|\chi\left(\{\vec{R}_I\};t\right)\right|^2 d\vec{R} = \langle\chi,t\,|\chi,t\rangle = 1, \quad \forall t\in\mathbb{R}, \tag{3.11}$$

where $d\vec{r} = d\vec{r}_1\cdots d\vec{r}_n$ and $d\vec{R} = d\vec{R}_1\cdots d\vec{R}_N$. The phase factor then can be written as

$$\tilde{E}_e = \int \phi^*\left(\{\vec{r}_i\};t\right)\chi^*\left(\{\vec{R}_i\};t\right)H_e\phi\left(\{\vec{r}_i\};t\right)\chi\left(\{\vec{R}_i\};t\right)d\vec{r}d\vec{R}. \tag{3.12}$$

By inserting (3.9) and (3.12) into (3.7), multiplying from the left by $\phi^*\chi^*$, integrating over all nuclear and electronic coordinates and imposing energy conservation condition

$$\frac{d}{dt}\int \Psi^*\left(\{\vec{r}_i\},\{\vec{R}_i\};t\right)H\Psi\left(\{\vec{r}_i\},\{\vec{R}_i\};t\right)d\vec{r}d\vec{R} = 0, \tag{3.13}$$

we obtain

$$i\hbar\frac{\partial}{\partial t}\phi = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2\phi + \left\{\int \chi^*\left(\{\vec{R}_I\};t\right)V_{n-e}\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)\chi\left(\{\vec{R}_I\};t\right)d\vec{R}\right\}\phi, \tag{3.14}$$

$$i\hbar\frac{\partial}{\partial t}\chi = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2\chi + \left\{\int \phi^*\left(\{\vec{r}_i\};t\right) H_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)\phi\left(\{\vec{r}_i\};t\right)d\vec{r}\right\}\chi, \quad (3.15)$$

which is a set of coupled time-dependent Schrödinger equations. The electrons move in a time-dependent effective potential given by averaging over nuclear degrees of freedom and vice versa.

Now we want to approximate the quantum mechanical equation for the nuclei (Eq. (3.15)) as a classical mechanical equation. To this aim, we rewrite the nuclear complex wavefunction $\chi\left(\{\vec{R}_I\};t\right)$ in a polar form

$$\chi\left(\{R_I\};t\right) = A\left(\{R_I\};t\right)\exp\left[iS\left(\{R_I\};t\right)/\hbar\right], \quad (3.16)$$

where A is the amplitude and S is the phase. Substituting Eq. (3.16) into Eq. (3.15) and separating the real and imaginary parts yields

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}(\nabla_I S)^2 + \int \phi^* H_e \phi d\vec{r} = \hbar^2 \sum_I \frac{1}{2M_I}\frac{\nabla_I^2 A}{A}, \quad (3.17)$$

$$\frac{\partial A}{\partial t} + \sum_I \frac{1}{M_I}(\nabla_I A)(\nabla_I S) + \sum_I \frac{1}{2M_I}A(\nabla_I^2 S) = 0 \quad (3.18)$$

Equation (3.18) for the amplitude A can be written as a continuity equation [257, 260, 261] by multiplying from the left by $2A$:

$$\frac{\partial A^2}{\partial t} + \sum_I \frac{1}{M_I}\nabla_I(A^2\nabla_I S) = 0 \quad (3.19)$$

Noting that the nuclear probability $\rho = |\chi|^2 = A^2$ and the current density $J_I = \frac{A^2(\nabla_I S)}{M_I}$, Eq. (3.19) becomes

$$\frac{\partial \rho}{\partial t} + \sum_I \nabla_I J_I = 0. \quad (3.20)$$

The continuity equation (3.20) ensures the conservation of particle probability χ^2 of the nuclei in the presence of a flux.

The expression of the phase S (Eq. (3.17)) contains one term that depends explicitly on the Planck's constant $\hbar$. This term vanishes in the classical limit $\hbar \to 0$, and Eq. (3.17) becomes

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}(\nabla_I S)^2 + \int \phi^* H_e \phi d\vec{r} = 0. \quad (3.21)$$

If we set $\nabla_I S$ to be the momentum of nucleus I

$$\vec{P}_I = \nabla_I S \quad (3.22)$$

then Eq. (3.21) becomes

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}\vec{P}_I^2 + \int \phi^* H_e \phi d\vec{r} = 0. \quad (3.23)$$

This equation is isomorphic to the equation of motion in the Hamiltonian-Jacobi formalism [257, 262] of classical mechanics

$$\frac{\partial S}{\partial t} + H\left(\{\vec{R}_I\}, \{\vec{P}_I\}\right) = 0, \tag{3.24}$$

in which

$$\begin{aligned} H\left(\{\vec{R}_I\}, \{\vec{P}_I\}\right) &= T\left(\{\vec{P}_I\}\right) + V\left(\{\vec{R}_I\}\right) \\ &= \sum_I \frac{1}{2M_I}\vec{P}_I^2 + \int d\vec{r}\phi^* H_e \phi, \end{aligned} \tag{3.25}$$

is the classical Hamilton function of the coordinates $\vec{R}_I$ and the conjugate momenta $\vec{P}_I$. The corresponding Newtonian equation of motion is

$$\frac{d\vec{P}_I}{dt} = -\vec{\nabla}_I \int d\vec{r}\phi^* H_e \phi = -\vec{\nabla}_I V_e\left(\{\vec{R}_I\}\right), \tag{3.26}$$

or

$$M_I \ddot{\vec{R}}_I = -\vec{\nabla}_I \int d\vec{r}\phi^* H_e \phi. \tag{3.27}$$

Therefore, in the classical limit ($\hbar \to 0$), the nuclei move as classical particles in an effective potential V_e, which is obtained by averaging over the electronic degrees of freedom at a given instantaneous nuclear positions $\vec{R}_I(t)$.

To calculate V_e we need to solve Eq. (3.14), which still contains the full quantum-mechanical nuclear wavefunction. Therefore, we need to apply the classical limit by replacing the nuclear density $\left|\chi\{\vec{R}_I\}\right|^2$ in Eq. (3.14) by a product of delta functions $\prod_I \delta\left[\vec{R}_I - \vec{R}_I(t)\right]$ centered at the instantaneous nuclear positions $R_I(t)$. The expectation value of the position operator becomes

$$\int d\vec{R}\chi^*\left(\{\vec{R}_I\}; t\right)\vec{R}_I\chi\left(\{\vec{R}_I\}; t\right) \to \vec{R}_I(t). \tag{3.28}$$

This leads to a time-dependent wave equation for the electrons

$$\begin{aligned} i\hbar\frac{\partial\phi}{\partial t} &= -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2\phi + V_{n-e}\left(\{\vec{r}_i\}; \{\vec{R}_I(t)\}\right)\phi \\ &= H_e\left(\{\vec{r}_i\}; \{\vec{R}_I(t)\}\right)\phi\left(\{\vec{r}_i\}; \{\vec{R}_I\}; t\right), \end{aligned} \tag{3.29}$$

which evolve self-consistently as the classical nuclei are propagated according to Eq. (3.27).

3.1.3 Ab initio molecular dynamics

3.1.3.1 Ehrenfest MD

In this *ab initio* MD scheme, the Newton's equations for the nuclei (Eq. (3.27)) are solved simultaneously with the time-dependent Schrödinger equation for the

electrons (Eq. (3.29)). The time-dependent Schrödinger equation of the electrons is solved "on-the-fly" as the nuclei are moving according to classical mechanics. This is the so-called Ehrenfest MD scheme. The total electronic wavefunction ϕ in Eq. (3.29) can be expanded in a basis of the electronic states ψ_l

$$\phi\left(\{\vec{r}_i\};\{\vec{R}_I\};t\right)=\sum_{l=0}^{\infty}c_l(t)\psi_l\left(\{\vec{r}_i\};\{\vec{R}_I\}\right), \tag{3.30}$$

where $c_l(t)$ are complex time-dependent coefficients, satisfying

$$\sum\nolimits_l |c_l(t)|^2=1. \tag{3.31}$$

The basis functions ψ_l can be chosen as the instantaneous eigenstates of the time-independent electronic Schrödinger equation

$$H_e\left(\{\vec{r}_i\};\{\vec{R}_I\}\right)\psi_l\left(\{\vec{r}_i\};\{\vec{R}_I\}\right)=E_l\left(\{\vec{R}_I\}\right)\psi_l\left(\{\vec{r}_i\};\{\vec{R}_I\}\right), \tag{3.32}$$

where $\vec{R}_I$ are the instantaneous nuclear positions at time t which are determined according to Eq. (3.27). With this choice of basis wavefunctions, the equations of motion (3.27) and (3.29) become

$$\begin{aligned} M_I\ddot{\vec{R}}_I(t) &= -\nabla_I\sum_k |c_k(t)|^2 E_k \\ &= -\sum_k |c_k(t)|^2\nabla_I E_k+\sum_{k,l}c_k^*c_l\left(E_k-E_l\right)\vec{d}_I^{kl}, \end{aligned} \tag{3.33}$$

$$i\hbar\dot{c}_k(t)=c_k(t)E_k-i\hbar\sum_l c_l(t)D^{kl}, \tag{3.34}$$

where

$$D^{kl}=\int\psi_k^*\frac{\partial}{\partial t}\psi_l d\vec{r}=\sum_I\dot{\vec{R}}_I\int\psi_k^*\nabla_I\psi_l d\vec{r}=\sum_I\vec{R}_I\vec{d}_I^{kl}. \tag{3.35}$$

The restriction to one electronic state (in most cases the ground state ψ_0) in Eq. (3.30) leads the equations of motion (3.27) and (3.29) to

$$M_I\ddot{\vec{R}}_I=-\vec{\nabla}_I\int d\vec{r}\psi_0^*H_e\psi_0, \tag{3.36}$$

$$i\hbar\frac{\partial\psi_0}{\partial t}=H_e\psi_0. \tag{3.37}$$

The propagation of the wavefunction in Eq. (3.37) is unitary. This implies that the wavefunction preserves its norm and the set of orbitals used to build up the wavefunction will stay orthonormal [257].

3.1.3.2 Born-Oppenheimer MD

Another approach to *ab initio* MD is to solve the time-independent Schrödinger equation for fixed instantaneous nuclear positions at each molecular dynamics step.

In this case, the intrinsic dynamics of the electrons is lost. The time-dependence of the electronic structure is imposed by the parametric dependence of H_e on the classical dynamics of the nuclei. The equations of motion of this so-called "Born-Oppenheimer" molecular dynamics are given by

$$M_I \ddot{\vec{R}}_I(t) = -\nabla_I \min_{\psi_0} \langle \psi_0 | H_e | \psi_0 \rangle , \tag{3.38}$$

$$H_e \psi_0 = E_0 \psi_0. \tag{3.39}$$

A drawback of the Born-Oppenheimer scheme is that the electronic wavefunction ψ_0 need to be obtained by minimizing $\langle H_e \rangle$ at every time step.

3.1.3.3 Car-Parrinello MD

In Ehrenfest MD method, the integration time step of the equations of motion is dictated by the intrinsic dynamics of the electrons, which is much faster than nuclear dynamics. Therefore, one has to use very short time step to integrate properly the equations of motion. In contrast, Born-Oppenheimer MD has no electronic dynamics. The electronic problem is treated through the time-independent Schrödinger equation. Only the time scale given by nuclear motion is relevant. This allows us to use a larger MD time step to integrate the equation of motion. However, Born-Oppenheimer MD requires that the electronic structure problem has to be solved at every MD step. This makes the method very expensive.

In 1985, Car and Parrinello [167] developed an *ab initio* MD scheme which combines advantages of both Ehrenfest and Born-Oppenheimer MD. This so-called Car-Parrinello MD allows the electronic wavefunction to be automatically minimized when the nuclei are propagated like in the Ehrenfest dynamics, but with an increased time step. Therefore, an explicit electronic minimization at each time step like in Born-Oppenheimer MD is avoided. The Car-Parrinello MD is based on the idea that fast electronic and slow nuclear motions can be adiabatically separated. Car and Parrinello introduced an extended Lagrangian which includes both electronic and nuclear degrees of freedom [167]. Each electronic degree of freedom (orbital) is considered as a "fictitious" particle with a fictitious mass μ. Given a sufficiently small value of the fictitious mass of the "electrons", the total electronic wavefunction can follow the nuclear motion adiabatically.

The Car-Parrinello Lagrangian is given by

$$L_{CP} = \sum_I \frac{1}{2} M_I \dot{\vec{R}}_I^2 + \sum_i \frac{1}{2} \mu \langle \dot{\varphi}_i | \dot{\varphi}_i \rangle - \langle \psi_0 | H_e | \psi_0 \rangle + \text{constraints}. \tag{3.40}$$

The first term is the nuclear kinetic energy, the second term is the kinetic energy of the fictitious electrons. μ is the fictitious mass. $\{\varphi_i\}$ is a set of orbitals used to expand the total electronic wavefunction ψ_0. $\langle \psi_0 | H_e | \psi_0 \rangle$ is the potential energy. The last term enforces the orthogonality of the orbitals φ_i, which is not conserved as in the case of the Ehrenfest scheme and therefore an orthonormalization is required at each step of the Car-Parrinello MD. The corresponding equations of motion for both nuclei and electrons are obtained from the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L_{CP}}{\partial \dot{\vec{R}}_I} = \frac{\partial L_{CP}}{\partial \vec{R}_I}, \tag{3.41}$$

$$\frac{d}{dt}\frac{\delta L_{CP}}{\delta \dot{\varphi}_i^*} = \frac{\delta L_{CP}}{\delta \varphi_i^*}, \tag{3.42}$$

which give, respectively,

$$M_I \ddot{\vec{R}}_I(t) = -\frac{\partial}{\partial \vec{R}_I}\langle\psi_0|\, H_e\, |\psi_0\rangle + \frac{\partial}{\partial \vec{R}_I}\{\text{constraints}\}, \tag{3.43}$$

$$\mu \ddot{\varphi}_i(t) = -\frac{\delta}{\delta \varphi_i^*}\langle\psi_0|\, H_e\, |\psi_0\rangle + \frac{\delta}{\delta \varphi_i^*}\{\text{constraints}\}. \tag{3.44}$$

The nuclei move at the physical temperature $\sim \sum_I M_I \dot{\vec{R}}_I^2$ whereas the electrons move at a "fictitious temperature" $\sim \sum_i \mu \langle \dot{\varphi}_i | \dot{\varphi}_i \rangle$. μ is the key parameter to control the adiabatic separation between the nuclei and the electrons. The smaller μ is the closer the instantaneous minimum energy $\langle\psi_0|\, H_e\, |\psi_0\rangle$ is to the exact Born-Oppenheimer energy surface and vice versa. On the other hand, small μ means small integration step. Therefore, in practice, a compromise between accuracy and efficiency determines the choice of μ.

3.1.3.4 Density Functional Theory

Any electronic structure method such as density functional theory (DFT), Hartree-Fock (HF), second-order many-body perturbation theory (MP2), or full configuration interaction (FCI) can, in principle, be used in combination with the *ab initio* molecular dynamics scheme. Here we focus only on the Hohenberg-Kohn-Sham density functional theory because the work in this thesis made extensive use of the DFT-based Car-Parrinello molecular dynamics.

DFT is founded on the two Hohenberg-Kohn (H-K) theorems. Consider a system of electrons moving under the influence of an external potential v_{n-e}.

(i) The first H-K theorem states that the external potential v_{n-e} is a unique functional of the ground-state electron density $\rho(\vec{r})$. Because v_{n-e} in turn determines the Hamiltonian H_e, full many-electron ground state is a unique functional of $\rho(\vec{r})$. The total ground-state energy E_0 can be written as a functional of $\rho(\vec{r})$

$$\begin{aligned} E_0[\rho(\vec{r})] &= T[\rho(\vec{r})] + E_{e-e}[\rho(\vec{r})] + E_{n-e}[\rho(\vec{r})] \\ &= F_{H-K}[\rho(\vec{r})] + \int \rho(\vec{r}) v_{n-e}(\vec{r}) d\vec{r}, \end{aligned} \tag{3.45}$$

where

$$v_{n-e}(\vec{r}) = -\frac{e}{4\pi\varepsilon_0}\sum_I \frac{Z_I}{\left|\vec{R}_I - \vec{r}\right|} + \frac{e^2}{4\pi\varepsilon_0}\sum_{I<J}\frac{Z_I Z_J}{\left|\vec{R}_I - \vec{R}_J\right|}. \tag{3.46}$$

$F_{H-K}[\rho(\vec{r})]$ is called the Hohenberg-Kohn functional. It is universal in the sense that it does not depend on the external potential v_{n-e}.

(ii) The second H-K theorem states that for a trial electron density $\tilde{\rho}(\vec{r})$, we always have

$$E_0 \leq E_0\left[\tilde{\rho}(\vec{r})\right]. \tag{3.47}$$

Therefore, if we want to find E_0, we need to search for the minimum of $E_0\left[\rho(\vec{r})\right]$.

Kohn-Sham approach

In order to find an explicit form for the functional $F_{H-K}\left[\rho(\vec{r})\right]$, Kohn and Sham assume that there is a reference system of non-interacting electrons (Kohn-Sham electrons described by the orbitals φ_i) that feels an effective potential (Kohn-Sham potential v_{ks}) such that its ground-state charge density is *idential* to that of the interacting system that we are dealing with.

The advantage of this auxiliary system is that for the non-interacting electrons, we know the exact solution to the equations of motion in terms of φ_i, which is the Slater determinant $\|\varphi_i\|$. The H-K functional F_{H-K} is then written in terms of the non-interacting system as

$$\begin{aligned} F_{H-K}\left[\rho(\vec{r})\right] = & \frac{-\hbar^2}{2m_e}\sum_i \int d\vec{r}\varphi_i^*(\vec{r})\nabla^2\varphi_i(\vec{r}) \\ & + \frac{e^2}{4\pi\varepsilon_0}\int d\vec{r}d\vec{r}'\rho(\vec{r})\frac{1}{|\vec{r}-\vec{r}'|}\rho(\vec{r}') + E_{\mathrm{xc}}\left[\rho(\vec{r})\right], \end{aligned} \tag{3.48}$$

where

$$\rho = \sum_i |\varphi_i|^2. \tag{3.49}$$

The total electronic energy is given by

$$\begin{aligned} E_{KS}\left[\rho(\vec{r})\right] &= F_{H-K}\left[\rho(\vec{r})\right] + \int d\vec{r}\rho(\vec{r})v_{n-e}(\vec{r}) \\ &= -\frac{\hbar^2}{2m_e}\sum_i \int d\vec{r}\varphi_i^*(\vec{r})\nabla^2\varphi_i(\vec{r}) + \frac{e^2}{4\pi\varepsilon_0}\int d\vec{r}d\vec{r}'\rho(\vec{r})\frac{1}{|\vec{r}-\vec{r}'|}\rho(\vec{r}') \\ &\quad + \int d\vec{r}\rho(\vec{r})v_{n-e}(\vec{r}) + E_{\mathrm{xc}}\left[\rho(\vec{r})\right], \end{aligned} \tag{3.50}$$

$E_{\mathrm{xc}}\left[\rho(\vec{r})\right]$ is called the exchange-correlation term, which is to be approximated. By applying the variational principle to (3.50) with the condition (3.49), we can derive a set of Kohn-Sham equations for the Kohn-Sham orbitals $\{\varphi_i\}$

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + v_{\mathrm{ks}}(\rho)\right]\varphi_i(\vec{r}) = \varepsilon_i\varphi_i(\vec{r}), \tag{3.51}$$

where $v_{\mathrm{ks}}(\rho)$ is the Kohn-Sham potential

$$v_{\mathrm{ks}}(\rho) = v_{\mathrm{n-e}}(\vec{r}) + \frac{1}{4\pi\varepsilon_0}\int \frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r}' + v_{\mathrm{xc}}(\rho), \tag{3.52}$$

and $v_{\mathrm{xc}}(\rho)$ is the exchange-correlation potential, which is defined as the functional derivative of E_{xc} with respect to $\rho(\vec{r})$

$$v_{\mathrm{xc}}\left[\rho(\vec{r})\right] = \frac{\delta E_{\mathrm{xc}}\left[\rho(\vec{r})\right]}{\delta\rho(\vec{r})}. \tag{3.53}$$

In terms of Kohn-Sham DFT, the Car-Parrinello Lagrangian is written as

$$L_{\mathrm{CP}} = \sum_I \frac{1}{2} M_I \dot{\vec{R}}_I^2 + \sum_i \frac{1}{2}\mu \int d\vec{r}\dot{\varphi}_i^*(\vec{r})\dot{\varphi}_i(\vec{r}) - E_{\mathrm{KS}}\left[\rho(\vec{r})\right] + \sum_{i,j}\Lambda_{ij}\left[\int d\vec{r}\varphi_i^*(\vec{r})\varphi_i(\vec{r}) - \delta_{ij}\right], \tag{3.54}$$

where Λ_{ij} are Lagrange multipliers.

Exchange-correlation functional

The exchange-correlation functional $E_{\mathrm{xc}}\left[\rho(\vec{r})\right]$ is unknown in the Kohn-Sham approach. Finding an approximation to this functional is critical to any application of DFT. The very first approximation to E_{xc} is the so-called "local density approximation" (LDA), which was introduced by Kohn and Sham [243, 244] and widely used in physics. In this approximation, the exchange-correlation energy at each point is assumed to be the same as that in a homogeneous electron gas with the same density

$$\begin{aligned} E_{\mathrm{xc}}^{\mathrm{LDA}}\left[\rho(\vec{r})\right] &= \int d\vec{r}\rho(\vec{r})\varepsilon_{\mathrm{xc}}^{\mathrm{hom}}\left[\rho(\vec{r})\right] \\ &= \int d\vec{r}\rho(\vec{r})\left\{\varepsilon_{\mathrm{x}}^{\mathrm{hom}}\left[\rho(\vec{r})\right] + \varepsilon_{\mathrm{c}}^{\mathrm{hom}}\left[\rho(\vec{r})\right]\right\}. \end{aligned} \tag{3.55}$$

For the homogeneous electron gas, the exchange part is known exactly [245, 246]

$$\varepsilon_{\mathrm{x}}\left[\rho(\vec{r})\right] = -\frac{3}{4}\left(\frac{3\rho(\vec{r})}{\pi}\right)^{\frac{1}{3}}. \tag{3.56}$$

No analytical expression is known for the correlation part ε_{c}. Based on highly accurate numerical quantum Monte-Carlo calculations of Cepperly *et al.* in 1980, many different analytical expressions of ε_{c} have been presented based on sophisticated interpolation schemes. The most widely used forms of ε_{c} are the one developed by Vosko, Wilk and Nusair [247] as well as the one given by Perdew and Wang [248].

For chemical applications, LDA is usually insufficient. This has led to the development of generalized-gradient approximations (GGAs), in which also the density gradient $\vec{\nabla}\rho(\vec{r})$ is taken into account in the expression of E_{xc}. In GGAs, the exchange-correlation functional is written as

$$\begin{aligned} E_{\mathrm{xc}}^{\mathrm{GGA}}\left[\rho(\vec{r})\right] &= \int d\vec{r}\rho(\vec{r})\varepsilon_{\mathrm{xc}}\left[\rho(\vec{r}), \left|\vec{\nabla}\rho(\vec{r})\right|\right] \\ &= E_{\mathrm{x}}^{\mathrm{GGA}}\left[\rho(\vec{r})\right] + E_{\mathrm{c}}^{\mathrm{GGA}}\left[\rho(\vec{r})\right] \end{aligned} \tag{3.57}$$

A very popular GGA functional is given by a combination of the Becke exchange functional and the Lee-Yang-Parr correlation functional (BLYP) [195, 196]. The Becke exchange functional is given by

$$E_{\mathrm{x}}^{\mathrm{GGA}}[\rho] = E_{\mathrm{x}}^{\mathrm{LDA}}[\rho] - \beta\int \rho(\vec{r})^{4/3}\frac{x^3}{1+6\beta\sinh^{-1}x}d\vec{r}, \tag{3.58}$$

where

$$x = \frac{\left|\vec{\nabla}\rho(\vec{r})\right|}{\rho(\vec{r})^{4/3}} \tag{3.59}$$

The parameter β is determined by a fit on the exact HF data and was fixed by Becke to be 0.0042 au.

The Lee, Yang and Parr correlation function was derived from the Colle-Salvetti formula to calculate the correlation energies from HF second order density matrix

$$E_{\mathrm{c}}^{\mathrm{GGA}}[\rho] = -a \int \frac{1}{1+\rho^{1/3}} \left(\rho + b\rho^{-2/3} \left[C_F \rho^{5/3} - 2t_W + \left(\frac{1}{9} t_W + \frac{1}{18} \nabla^2 \rho \right) \right] e^{-\rho^{1/3}} \right) d\vec{r}, \tag{3.60}$$

where

$$C_F = \frac{3}{10} \left(3\pi^2\right)^{2/3}, \tag{3.61}$$

$$t_W = \frac{1}{8} \frac{|\nabla\rho|^2}{\rho} - \frac{1}{8} \nabla^2 \rho, \tag{3.62}$$

and $a = 0.04918$, $b = 0.132$.

3.1.3.5 Basis set

In practical implementations for solving the Kohn-Sham equations (3.51), the Kohn-Sham orbitals $\{\varphi_i(\vec{r})\}$ are expanded as a linear combination of basis functions with well-known behaviors and with which it is possible to perform mathematical operations very fast on a computer

$$\varphi_i(\vec{r}) = \sum_l c_{il} f_l \left(\vec{r}; \left\{ \vec{R}_I \right\} \right). \tag{3.63}$$

Localized basis sets

In quantum chemistry, the two most frequently used basis functions are the Slater-type basis functions

$$f_l^S = N_l^S r_x^{mx} r_y^{my} r_z^{mz} \exp\left[-\zeta_m \left|\vec{r}\right|\right], \tag{3.64}$$

and the Gaussian-type basis functions

$$f_l^G = N_l^G r_x^{mx} r_y^{my} r_z^{mz} \exp\left[-\alpha_m r^2\right], \tag{3.65}$$

where N_m, ζ_m and α_m are constants and kept fixed during an electronic structure calculation, and only the coefficients c_{il} need to be optimized.

Plane wave basis set and pseudopotentials

In periodic systems, it is more convenient to use plane waves as a basis set. Plane wave functions are defined as

$$f_G^{PW} = \frac{1}{\sqrt{V}} e^{i\vec{G}\vec{r}}, \tag{3.66}$$

where V is the volume of the periodic cell, $\vec{G}$ is the reciprocal lattice vector.

According to Bloch's theorem, the eigenfunctions of periodic systems can be written as

$$\varphi_{i,\vec{k}}(\vec{r}) = e^{i\vec{k}\vec{r}} u_{i,\vec{k}}(\vec{r}), \tag{3.67}$$

where $\vec{k}$ is a wave vector in the first Brillouin zone of the reciprocal lattice. $u_{i,\vec{k}}(\vec{r})$ is a cell-periodic function

$$u_{i,\vec{k}}(\vec{r}+\vec{R}) = u_{i,\vec{k}}(\vec{r}), \tag{3.68}$$

which can be expanded as a Fourier series

$$u_{i,\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} \sum_{\vec{G}} C_{i,\vec{k}}(\vec{G}) e^{i\vec{G}\vec{r}}. \tag{3.69}$$

Eq. (3.67) then becomes

$$\varphi_{i,\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} \sum_{\vec{G}} c_{i,\vec{k}}(\vec{G}) e^{i(\vec{G}+\vec{k})\vec{r}}. \tag{3.70}$$

A further approximation can be introduced to minimize the size of the plane wave basis set: the so-called pseudopotential approximation [249, 250, 251]. This approximation is based on the observation that core electrons, *i.e.* the electrons close to the nuclei that are not involved in the chemical reactions, are somewhat unaffected by the chemical environment. Therefore, it is sufficient to consider just the valence electrons, whereas the core ones are replaced by pseudopotentials to take into account their effect in the motion of the valence electron. The valence electron wavefunction can then be replaced by pseudo wavefunction $|\psi^{\mathrm{ps}}\rangle$, which now varies smoothly in the core region (Fig. 3.1). As a consequence, the number of plane waves needed to expand the wavefunctions is greatly reduced.

The pseudo valence states can be found by solving the all-electron Schrödinger equation for an atom

$$\left[\hat{H} + \hat{V}_{\mathrm{nl}}\right] |\psi^{\mathrm{ps}}\rangle = E\, |\psi^{\mathrm{ps}}\rangle\,, \tag{3.71}$$

where $\hat{H}$ is all-electron Hamiltonian of an atom, $\hat{V}_{\mathrm{nl}}$ is an energy-dependent non-local potential

$$\hat{V}_{\mathrm{nl}} = \sum_{i \in \mathrm{core}} (E - E_i)\, |\chi_i\rangle\, \langle\chi_i.| \tag{3.72}$$

$|\chi_i\rangle$ and E_i are eigenstates and eigenvalues of the core-electron states, respectively. Since the additional potential $\hat{V}_{\mathrm{nl}}$ is repulsive, sum of $\hat{V}_{\mathrm{nl}}$ with the strongly attractive Coulomb potential $\hat{V}_{\mathrm{Cou}}$ results in a much weaker pseudopotential [251, 252]

$$\hat{V}^{\mathrm{PS}} = \hat{V}_{\mathrm{Cou}} + \sum_{i \in \mathrm{core}} (E - E_i)\, |\chi_i\rangle\, \langle\chi_i.|. \tag{3.73}$$

Pseudopotentials are constructed from *ab initio* calculations for isolated atoms, *e.g.*, by solving the Kohn-Sham equations. For atoms, due to the spherical symmetry, the wavefunctions can be written as a product of a radial function and a spherical harmonic. The Kohn-Sham equations then reduce to one-dimensional differential equations for the radial function. Serveral schemes for generating pseudopotentials have been proposed. In this thesis, we have used the "Norm-conserving" pseudopotentials derived from the Martins-Troullier (MT) scheme [197], in which the pseudopotentials have to satisfy the following conditions:

(i) The lowest pseudo wavefunction should not contain any node.

(ii) The eigenvalues of both real and pseudo wavefunctions must be the same

(iii) The radial pseudo wavefunction with an angular momentum l should be equal to the radial all-electron wavefunction outside a cut-off radius r_c.

(iv) The integrated charge inside r_c must be the same for both wavefunctions.

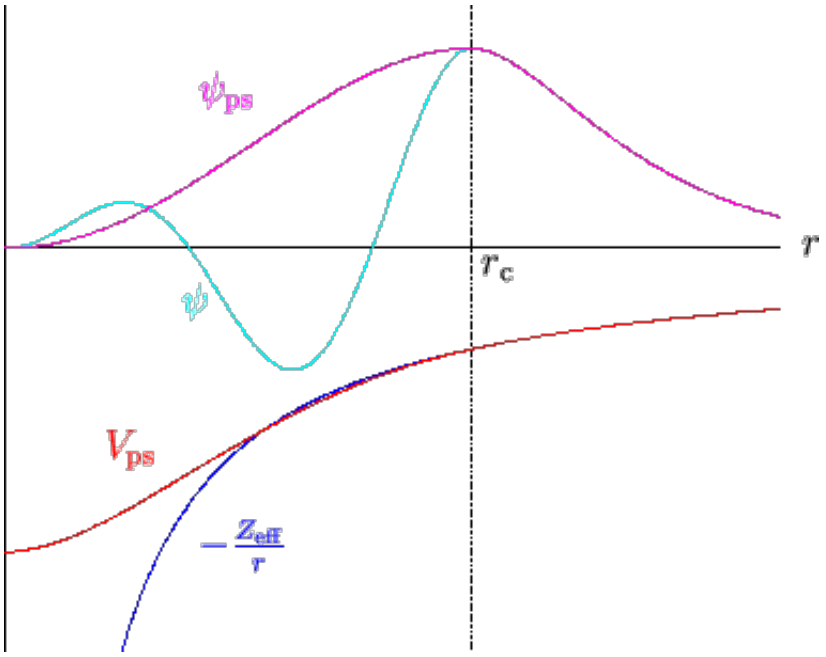


Figure 3.1 Schematic diagram of the relationship between all-electron and pseudopotentials and wavefunctions as well as real and pseudopotentials, adapted from [253].

3.1.4 Force field-based molecular dynamics

Due to a large number of atoms (thousands of atoms) in a typical biomolecular systems, full *ab initio* molecular dynamics is not feasible. Therefore, force fields are usually used to approximate the interatomic potential energy. In this approach, the potential energy of the atoms is approximated by a simple analytical function (force field) of the atomic coordinates. In this way, the need to solve the quantum electronic problem can be avoided.

Force fields contain a set of adjustable parameters, which have to be fitted to data obtained from experiments or high-level quantum mechanical calculations. There are various types of force fields used in biomolecular simulations with different functional forms and parameter sets, such as AMBER [173], GROMOS [189], CHARMM [254], etc.

The AMBER force field [173] which has been used in this work has the following functional form

$$U^{\mathrm{FF}} = \sum_{\text{bonds}} K_R \left(R - R_{\mathrm{eq}}\right)^2 + \sum_{\text{angles}} K_\theta \left(\theta - \theta_{\mathrm{eq}}\right)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2} \left[1 + \cos(n\phi - \gamma)\right] + \sum_{I<J} \frac{q_I q_J}{\varepsilon R_{IJ}} + \sum_{I<J} \sigma_{IJ} \left[\left(\frac{R_{\min I,J}}{R_{IJ}}\right)^{12} - 2\left(\frac{R_{\min I,J}}{R_{IJ}}\right)^{6}\right]. \tag{3.74}$$

K_R, K_θ and V_n are bond, angle and dihedral force constants, respectively. R and R_{eq} are bond length and its equilibrium value, respectively. θ and θ_{eq} are the angle and its equilibrium value, respectively. n, ϕ and γ are the number of barriers, dihedral angle and phase, respectively. R_{IJ} is the distance between atom I and atom J. q_I is the partial charge of atom I. ε is the dielectric constant. $R_{\min I,J}$ and σ_{IJ} are the equilibrium distance and the well depth of the van der Waals interaction potential, respectively.

3.1.5 Hybrid Car-Parrinello/molecular mechanics (QM/MM) molecular dynamics

QM/MM MD is the method of choice for studies of chemical reactions and for accurate descriptions of metallic complexes in large biomolecular systems [172, 170, 171]. In this approach, the system of interest is partitioned into two parts: one is treated at quantum mechanical level (QM part), and the other one treated with the molecular mechanics force fields (the MM part).

The Car-Parrinello MD scheme (Eq. 3.54) can also be extended to deal with QM/MM approach. The corresponding extended Lagrangian is [169]

$$\begin{aligned} L_{\mathrm{QM/MM}} &= L_{\mathrm{CP}} - E_{\mathrm{MM}} - E_{\mathrm{QM/MM}} \\ &= \frac{1}{2}\sum_I M_I \dot{\vec{R}}_I^2 + \frac{1}{2}\mu \sum_i \int d\vec{r}\, \dot{\varphi}_i^*(\vec{r})\dot{\varphi}_i(\vec{r}) - E_{\mathrm{KS}}\left[\rho(\vec{r})\right] \\ &\quad + \sum_{ij} \Lambda_{ij} \left[\int d\vec{r}\, \varphi_i^*(\vec{r})\varphi_i(\vec{r}) - \delta_{ij}\right] - E_{\mathrm{MM}} - E_{\mathrm{QM/MM}}, \end{aligned} \tag{3.75}$$

where E_{MM} is the force field-based energy function describing the MM part, $E_{\mathrm{QM/MM}}$ is the coupling term between QM and MM parts.

In our work, E_{MM} is given by the AMBER force field in Eq. (3.74)

$$E_{\mathrm{MM}} = U^{\mathrm{FF}} \tag{3.76}$$

The QM/MM coupling term $E_{\mathrm{QM/MM}}$ can be expanded in three terms:

$$E_{\mathrm{QM/MM}} = E_{\mathrm{QM/MM}}^{\mathrm{ele}} + E_{\mathrm{QM/MM}}^{\mathrm{vdw}} + E_{\mathrm{QM/MM}}^{\mathrm{bonded}}. \tag{3.77}$$

The first term in the right-hand side is the electrostatic coupling between the quantum charge distribution and the classical point charges. The second and third terms describe the van der Waals and bonded interactions between QM and MM atoms,

respectively. Covalent bonds cut by the QM/MM boundary are saturated by capping hydrogen atoms. The bonded interactions between QM and MM regions, *i.e.* bond stretching, angle bending and dihedral torsions are described by the force field. The same holds for van der Waals interaction between QM and MM atoms.

The electrostatic coupling term $E^{\text{ele}}_{\text{QM/MM}}$ reads

$$E^{\text{ele}}_{\text{QM/MM}} = \sum_{i \in MM} q_i \int d\vec{r} \rho(\vec{r}) v_i \left(|\vec{r} - \vec{r}_i| \right) \tag{3.78}$$

where q_i are the partial charges at $\vec{r}_i$ and $v_i \left(|\vec{r} - \vec{r}_i| \right)$ is a modified Coulomb potential. This potential is modified at short range to avoid the so-called electron spill-out phenomenon, *i.e.* the unphysical accumulation of charge density at the boundary of the QM box due to the presence of classical positive charges of the nearby MM region [169]

$$v_i(\vec{r}) = \frac{r_{ci}^n - r^n}{r_{ci}^{n+1} - r^{n+1}}. \tag{3.79}$$

In Eq. (3.79), n is usually fixed to 4 and r_{ci} is the covalent radius of atoms. $v_i(r)$ is chosen such that it saturates to $1/r$ at large distance and goes smoothly to a constant value for small r (see Fig. 3.2).

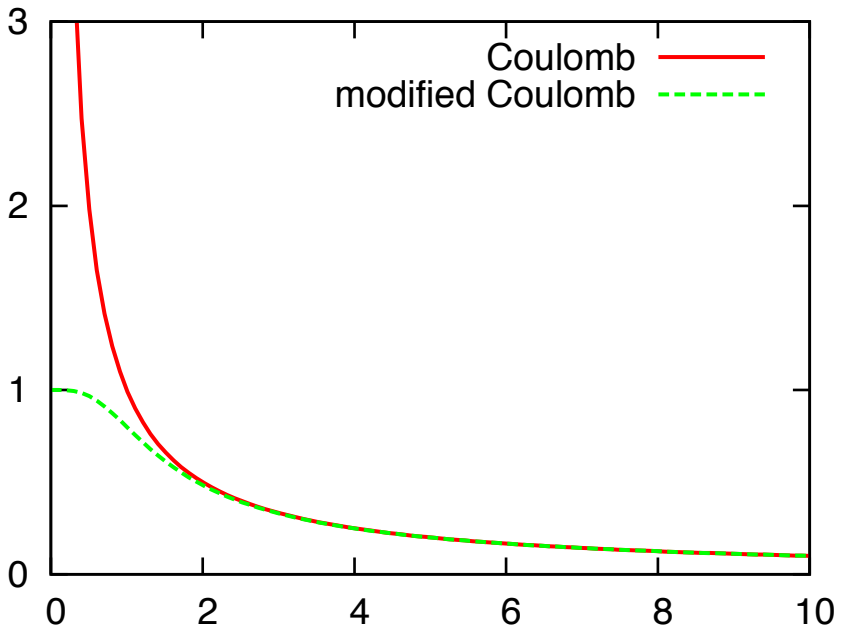


Figure 3.2 Schematic representation of Coulomb and modified Coulomb potentials.

The influence of r_c's on the pair correlation function between QM and MM atoms was probed by performing a QM/MM simulations of a quantum water molecule in a box of classical water [169]. It was shown that the QM/MM pair correlation function were consistent with the full quantum calculation when r_{ci} is chosen close to the covalent radius (*i.e.* ~ 0.4 Å for hydrogen and ~ 0.7 Å for oxygen).

Because the calculation of the integrals in Eq. (3.78) is computationally expensive, they are calculated only for the MM atoms within a cut-off radius (~ 5 Å) from

the interface. Beyond this first shell and within a second cut-off radius ($\sim$ 10 Å), the electrostatic interactions are calculated using D-RESP (Dynamically Generated Potential Derived Charge) charge scheme [168]. In this approach, a set or point charges for QM atoms are fitted "on the fly" in each step of QM/MM simulations so as to reproduce the electrostatic potential evaluated at the positions of MM atoms close to the QM region. These point charges then interact with the point charges of MM atoms within the second shell. Finally, the electrostatic interactions are calculated between a multipolar expansion of the QM charge density and the MM classical point charges contained in the outer most region ($>$ 10 Å) [169].

3.1.6 Force matching procedure

Although QM/MM method can provide very accurate description of metal coordination structures in large biomolecules, the typical time scales are rather short ($\leq$ 100 ps) with respect to the time scales of relevant biological phenomena. To study global conformational changes of large systems, extensive classical MD simulations are, therefore, mandatory. This issue can be addressed by developing a set of force field parameters for the metal moieties in the biomolecules. Then, long time classical MD simulations can be performed based on these newly derived parameters.

The force matching procedure [174, 175] based on QM/MM trajectories offers an accurate and elegant method for the force field parameterization. In this approach, a QM/MM simulation is performed for a given system, which is divided into two parts: one treated at the quantum mechanical level (the QM part) and the other at the force field level (the MM part). The fragment that needs to be parameterized is included in the QM part. A set of snapshots is selected along the QM/MM trajectory. This set is used as reference structures for the force matching procedure. The *ab initio* forces acting on the QM atoms together with the electrostatic potential and field at positions of the nearby MM atoms are stored for each selected conformation.

First a set of point charges $\{q_I\}$ for the atoms in the QM region is derived. This set is fitted to reproduce both the electrostatic potential and the electrostatic field that the QM part exerts on the surrounding MM atoms. The following target function is minimized to obtain $\{q_I\}$

$$\chi^2 = \sum_{l}^{L} \left(\sum_{J \in NN} w^v (V_{Jl}^{MM} - V_{Jl}^{\rho})^2 + \sum_{J \in NN} w^E \left\| \vec{E}_{Jl}^{MM} - \vec{E}_{Jl}^{\rho} \right\|^2 + \sum_{I \in QM} w^H (q_I - q_{Il}^H)^2 \right) + w^Q \left(Q^{\text{tot}} - \sum_{I \in QM} q_I \right)^2 . \tag{3.80}$$

The index l runs over L conformations. The index J runs over all MM atoms that interact explicitly with the quantum electron density in the QM/MM calculations (here called NN atoms). The index I runs over all QM atoms. V_{Jl}^{ρ} and $\vec{E}_{Jl}^{\rho}$ are the electrostatic potential and the electrostatic field, respectively, due to the presence of the QM charge density, and felt by classical atom J in configuration l. V_{Jl}^{MM} and $\vec{E}_{Jl}^{MM}$ are the potential and field generated by the $\{q_I\}$ set. Restraining the $\{q_I\}$ set so as to reproduce both the electrostatic potential and the electrostatic field has been shown to provide better results than reproducing only the electrostatic

potential [174]. The third term restrains the set of charges $\{q_I\}$ to the Hirshfeld charge values q_{Il}^H [272]. The last term restrains the total charge of the $\{q_I\}$ set to the correct value Q^{tot}.

Then, the $\{q_I\}$ and the van der Waals parameters (taken from a standard force field such as AMBER [173]) are used to calculate the total non-bonded forces acting on the QM atoms. These forces are subtracted from the total force, yielding a set of bonded forces

$$\vec{F}^{QM_{\text{bonded}}} = \vec{F}^{QM} - \vec{F}^{MM_{\text{non-bonded}}}. \tag{3.81}$$

Finally, the force field bonded parameters, including bonds, angles, and torsions for the atoms in the QM region are obtained through a least squares fit to the bonded *ab initio* forces.

$$\sigma^2 = \sum_{l=1}^{L} \sum_{I \in QM} \left\| \vec{F}_{lI}^{MM_{\text{bonded}}} - \vec{F}_{lI}^{QM_{\text{bonded}}} \right\|^2. \tag{3.82}$$

The parameters are fitted to obtain the smallest average error over all the reference frames, extracted from QM/MM simulations at finite temperature.

3.1.7 Ewald summation

In molecular dynamics simulations of biomolecules, the computation of non-bonded interactions, *e.g.*, the fourth and fifth terms in Eq. (3.74) is the most expensive because the sums are over all the pairs of atoms. Furthermore, under the periodic boundary conditions, these sums extend over an infinite number of periodic images. The simplest way to avoid summing over all the atom pairs is to truncate the potential at a cut-off distance R_c. In such a case, one can show that [256] the contribution from the tail of the potential $U(R)$ can be estimated (in 3D) using

$$U^{\text{tail}} = \frac{Nn}{2} \int_{R_c}^{\infty} 4\pi R^2 U(R) dR, \tag{3.83}$$

where n is the average number density. The equation above implies that the tail contribution to the potential energy diverges, unless the potential energy $U(R)$ decays faster than R^{-3}. Therefore, we can safely apply the truncation method for the van der Waals term (the fifth term in Eq. (3.74)) which decays as R^{-6}, but not for the Coulombic term (the fourth term in Eq. (3.74)) which decays as R^{-1}.

The Ewald summation method offers a fast way to compute the long-range Coulombic interaction in periodic systems as the fourth term in Eq. (3.74). Let us suppose that the system consists of N charged particles in a cubic box of size L, subject to periodic boundary conditions. The total Coulomb energy of the system is

$$U = \frac{1}{2} \sum_{\vec{n}}{}' \sum_{IJ} \frac{q_I q_J}{\left| \vec{R}_{IJ} + \vec{n} \right|}, \tag{3.84}$$

where q_I is the charge of particle I. The prime on the summation indicates the sum is over all periodic images $\vec{n}$ and over all particles I, J, except $I = J$ if $\vec{n} = 0$. The above sum extends over an infinite number of periodic images. Moreover, it

is conditionally convergent, which means that the result depends on the order of summation [255]. Therefore, it cannot be used to compute electrostatic energy in molecular dynamics simulations.

In the Ewald summation method [263, 264], the charge distribution $\rho(\vec{R})$ of the system is represented by an infinite set of point charges q_I, which can be written in terms of delta function as

$$\rho_I\left(\vec{R}\right) = q_I\delta\left(\vec{R} - \vec{R}_I\right). \tag{3.85}$$

At each point charge, an isotropic Gaussian charge distribution of equal magnitude but opposite sign:

$$\rho_I^G\left(\vec{R}\right) = q_I\left(\frac{2}{\sqrt{\pi}}\right)^3 e^{-\alpha^2\left|\vec{R}-\vec{R}_I\right|^2} \tag{3.86}$$

is then added. This smeared charge screens the interaction between the point charges, so that the interaction of the screened charge distribution

$$\rho_I^S\left(\vec{R}\right) = \rho_I\left(\vec{R}\right) + \rho_I^G\left(\vec{R}\right) \tag{3.87}$$

becomes short range:

$$U^S = \frac{1}{2}\sum_{\vec{n}}\sum_{IJ}\frac{q_Iq_J}{\left|\vec{R}_{IJ}+\vec{n}\right|}\mathrm{erfc}\left(\alpha\left|\vec{R}_{IJ}+\vec{n}\right|\right), \tag{3.88}$$

where

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-y^2}dy, \tag{3.89}$$

is the complementary error function. Therefore, U^S can be approximated by truncation.

Finally, a Gaussian charge distribution of opposite sign $-\rho_I^G\left(\vec{R}\right)$ must be added to recover the original charge distribution $\rho_I\left(\vec{R}\right)$. The interaction energy of this Gaussian distribution can be expressed as a lattice sum in the reciprocal space minus a self term:

$$U^m = \frac{1}{2\pi V}\sum_{IJ}q_Iq_J\sum_{\vec{k}\neq 0}\frac{1}{k^2}\left(e^{-\frac{k^2}{4\alpha^2}}e^{i\vec{k}\vec{R}_{IJ}}\right) - \frac{\alpha}{\sqrt{\pi}}\sum_I q_I^2, \tag{3.90}$$

where

$$\vec{k} = 2\pi L^{-1}\left(l_x, l_y, l_z\right), \tag{3.91}$$

and l_x, l_y and l_z are integers. Due to the presence of the exponential factor the lattice sum can be well approximated by a finite one.

3.1.8 Molecular dynamics in NPT ensemble

The trajectory generated by integrating the Newton's second equation (3.27) corresponds to a sampling on an NVE ensemble. However, most experiments are performed at constant pressure and constant temperature conditions. Therefore, it is essential to be able to perform simulations under the same NPT condition. This requires introducing thermostat and barostat algorithms to control the temperature and pressure, respectively.

3.1.8.1 Thermostats

Berendsen thermostat

Berendsen thermostat [265] is a weak coupling method, in which the temperature of the system is kept close to a target by the equation

$$\frac{dT}{dt} = \frac{1}{\tau_T}\left[T_0 - T(t)\right], \tag{3.92}$$

where T is the instantaneous value of the temperature and τ_T is the coupling time. The instantaneous temperature T of the system with $3N$ degrees of freedom is related to the kinetic energy through the equation

$$E_{\text{kin}} = \sum_I M_I \dot{\vec{R}}_I^2(t) = \frac{3}{2} N k_B T(t). \tag{3.93}$$

Therefore, to change the temperature by an amount δT at each finite time step δt we have to change kinetic energy by an amount

$$\delta E_{\text{kin}} = \frac{3}{2} N k_B \delta T = \frac{3}{2} N k_B \frac{1}{\tau_T}\left[T - T(t)\right]\delta t. \tag{3.94}$$

This change in kinetic energy can be achieved by rescaling the velocities by a factor λ, which is related to δE_{kin} as

$$\delta E_{\text{kin}} = \sum_I M_I \left(\lambda^2 - 1\right) \dot{\vec{R}}_I^2(t). \tag{3.95}$$

From Eqs. (3.93), (3.94) and (3.95) we obtain

$$\lambda = \sqrt{1 + \frac{\delta t}{\tau_T}\left(\frac{T_0}{T} - 1\right)}. \tag{3.96}$$

The factor λ is used to scale the velocities at each integration step to relax the temperature towards the target value T_0. The relaxation rate is controlled by the coupling constant τ_T.

Nose-Hoove thermostat The approach of Nose [199] is based on the use of extended Lagrangian. Nose introduce an extra degree of freedom S to the Lagrangian of the classical N-body system

$$L_{\text{Nose}} = \frac{1}{2}\sum_{I=1}^{N} M_I S^2 \dot{\vec{R}}_I^2 - U\left(\left\{\vec{R}_I\right\}\right) + \frac{Q}{2}\dot{S}^2 - \frac{L}{\beta}\ln S, \tag{3.97}$$

where L is a constant, Q is an effective mass associated with S. The conjugate momenta of $\vec{R}_I$ and S are given by

$$\vec{P}_I \equiv \frac{\partial L_{\text{Nose}}}{\partial \dot{\vec{R}}_I} = M_I S^2 \dot{\vec{R}}_I \tag{3.98}$$

$$\vec{P}_S \equiv \frac{\partial L_{\text{Nose}}}{\partial \dot{S}} = Q\dot{S} \tag{3.99}$$

Hence, the Hamiltonian of the extended system of N particles and the additional degree of freedom S is

$$H_{\text{Nose}} = \sum_I \frac{\vec{P}_I^2}{2M_I S^2} + U\left(\left\{\vec{R}_I\right\}\right) + \frac{P_S^2}{2Q} + L\frac{\ln S}{\beta}. \tag{3.100}$$

The corresponding partition function of the microcanonical (NVE) ensemble is

$$Z_{\text{Nose}} = \frac{1}{N!}\int dP_S dS d\vec{P} d\vec{R}\delta\left(E - H_{\text{Nose}}\right). \tag{3.101}$$

Let us define the scaled momenta

$$\vec{P}' = \frac{\vec{P}}{S}. \tag{3.102}$$

Z_{Nose} becomes

$$Z_{\text{Nose}} = C\frac{1}{N!}\int d\vec{P}' d\vec{R}\exp\left[-\beta\frac{3N+1}{L}H\left(\vec{P}', \vec{R}\right)\right], \tag{3.103}$$

where

$$H\left(\vec{P}', \vec{R}\right) = \sum_I \frac{\vec{P}_I'^2}{2M_I} + U\left(\left\{\vec{R}_I\right\}\right). \tag{3.104}$$

If we choose $L = 3N + 1$, then Z_{Nose} corresponds to the partition function of the canonical (NVT) ensemble of the system with the degrees of freedom $\left\{\vec{P}_I'\right\}, \left\{\vec{R}_I\right\}$. Therefore, we can interpret $\vec{P}'$ as the physical momenta while $\vec{P}$ as the virtual momenta. If we denote all the physical variables as primed and virtual momenta as unprimed, we have the following relations

$$\begin{aligned} \vec{R}' &= \vec{R}, \\ \vec{P}' &= \frac{\vec{P}}{S}, \\ S' &= S, \\ \Delta t' &= \frac{\Delta t}{S}. \end{aligned} \tag{3.105}$$

From the Hamiltonian (3.100), one can derive the equations of motion for the virtual variables

$$\frac{d\vec{R}_I}{dt} = \frac{\partial H_{Nose}}{\partial \vec{P}_I} = \frac{\vec{P}_I}{M_I S^2}, \tag{3.106}$$

$$\frac{d\vec{P}_I}{dt} = -\frac{\partial H_{\text{Nose}}}{\partial \vec{R}_I} = -\frac{\partial U\left(\left\{\vec{R}_I\right\}\right)}{\partial \vec{R}_I}, \tag{3.107}$$

$$\frac{dS}{dt} = \frac{\partial H_{\text{Nose}}}{\partial P_S} = \frac{P_S}{Q}, \tag{3.108}$$

$$\frac{dP_S}{dt} = -\frac{\partial H_{\text{Nose}}}{\partial S} = \left(\sum_I \frac{\vec{P}_I^2}{M_I S^2} - \frac{L}{\beta}\right)/S, \tag{3.109}$$

or for the physical variables, these equations of motion become

$$\frac{d\vec{R}'_I}{dt'} = \frac{\vec{P}'_I}{M_I}, \tag{3.110}$$

$$\frac{d\vec{P}'_I}{dt'} = -\frac{\partial U\left(\left\{\vec{R}'_I\right\}\right)}{\partial \vec{R}'_I} - \left(\frac{S'P'_S}{Q}\right)\vec{P}'_I, \tag{3.111}$$

$$\frac{1}{S}\frac{dS'}{dt'} = \frac{S'P'_S}{Q}, \tag{3.112}$$

$$\frac{d\left(\frac{S'P'_S}{Q}\right)}{dt'} = \left(\sum_I \frac{\vec{P'}_I^{\,2}}{M_I} - \frac{L}{\beta}\right)/Q. \tag{3.113}$$

Hoover [200] further simplified these equations by introducing the thermodynamic friction coefficient

$$\xi = \frac{S'P'_S}{Q}. \tag{3.114}$$

The equations of motions (3.110)–(3.113) become, after dropping the primes

$$\dot{\vec{R}}_I = \frac{\vec{P}_I}{M_I}, \tag{3.115}$$

$$\dot{\vec{P}}_I = \frac{\partial U\left(\left\{\vec{R}_I\right\}\right)}{\partial \vec{R}_I} - \xi\vec{P}_I, \tag{3.116}$$

$$\frac{\dot{S}}{S} = \frac{d\ln S}{dt} = \xi. \tag{3.117}$$

$$\dot{\xi} = \left(\sum_I \frac{\vec{P}_I^2}{M_I} - \frac{L}{\beta}\right)/Q, \tag{3.118}$$

3.1.8.2 Barostats

In a molecular dynamics simulation of a periodic system in a box, the isotropic pressure p is defined as the trace of the pressure tensor $\mathbf{p}$

$$p = \frac{1}{3}\mathrm{Tr}\left(\mathbf{p}\right). \tag{3.119}$$

$\mathbf{p}$ is calculated by the Clausius virial theorem

$$\mathbf{p} = \frac{2}{V}\left(E_{\mathrm{Kin}} - \Xi\right), \tag{3.120}$$

where V is the box volume, E_{Kin} is the kinetic energy and Ξ is the inner virial tensor

$$\Xi = -\frac{1}{2}\sum_{I<J}\vec{R}_{IJ}\vec{F}_{IJ}. \tag{3.121}$$

Pressure control during a simulation can be achieved through a change in the inner virial tensor by rescaling the inter-particle distances R_{IJ}.

In the Andersen barostat approach [266], the coordinates $\left\{\vec{R}_I\right\}$ are replaced by scaled coordinates $\left\{\vec{R}'_I\right\}$ defined as

$$\vec{R}'_I = \frac{\vec{R}_I}{V^{1/3}}. \tag{3.122}$$

Consider a new Lagrangian, in which the volume V appears as an additional dynamical variable

$$L_{\mathrm{An}} = \frac{1}{2}V^{2/3}\sum_I M_I \dot{\vec{R}}'^2_I - \sum_{I<J} U\left(V^{1/3}R'_{IJ}\right) + \frac{1}{2}\mu\dot{V}^2 - p_0 V. \tag{3.123}$$

The first two terms on the right hand side are the Lagrangian of the unscaled system, the third term is the kinetic energy for the motion of V, and the fourth term represents a potential energy associated with V. p_0 and μ are constants, whose physical meaning is represented by the target pressure and the pressure coupling constant, respectively.

From the Lagrangian (3.123), the corresponding Hamiltonian and hence the equations of motion for the scaled variables can be derived. Finally, mapping back to the unscaled variables $\left\{\vec{R}_I\right\}$, we get

$$\frac{d\vec{R}_I}{dt} = \frac{\vec{P}_I}{M_I} + \frac{1}{3}\vec{R}_I\frac{d\ln V}{dt}, \tag{3.124}$$

$$\frac{d\vec{P}_I}{dt} = -\sum_{I<J}\vec{R}_{IJ}U'\left(R_{IJ}\right) - \frac{1}{3}\vec{P}_I\frac{d\ln V}{dt}, \tag{3.125}$$

$$\mu\frac{d^2V}{dt^2} = p_0 + \left[\frac{2}{3}\sum_I\frac{\vec{P}_I^2}{2M_I} - \frac{1}{3}\sum_{I<J}\vec{R}_{IJ}U'\left(R_{IJ}\right)\right]/V. \tag{3.126}$$

Parrinello and Rahman [202] in 1981 extended the Andersen barostat to allow the simulation box to change its shape. The cell can have an arbitrary shape and its volume is completely determined by three vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$

$$V = \det\mathbf{h} = \vec{a}\cdot\left(\vec{b}\times\vec{c}\right), \tag{3.127}$$

where $\mathbf{h}$ is a 3 matrix, whose columns are the three vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$.

The position $\vec{R}_I$ of a particle can be written in terms of $\mathbf{h}$ and a column vector $\vec{S}_I$ with components ξ_I, η_I and ζ_I as

$$\vec{R}_I = \mathbf{h}\vec{S}_I = \xi_I\vec{a} + \eta_I\vec{b} + \zeta_I\vec{c}, \tag{3.128}$$

with $0 \leq \xi_I, \eta_I, \zeta_I \leq 1$. The squared distance between particle I and J is given by

$$R^2_{IJ} = S_I^T\mathbf{G}S_J, \tag{3.129}$$

where $\mathbf{G}$ is the metric tensor

$$\mathbf{G} = \mathbf{h}^T\mathbf{h}. \tag{3.130}$$

Using notations above, the Lagrangian can then be written as

$$L_{\mathrm{PR}} = \frac{1}{2}\sum_I M_I \dot{S}_I^T \mathbf{G}\dot{S}_I - \sum_{I<J} U\left(R_{IJ}\right) + \frac{1}{2}\mu \mathrm{Tr}\left(\dot{\mathbf{h}}^T\dot{\mathbf{h}}\right) - p_0 V. \tag{3.131}$$

The corresponding equations of motion can be derived similarly to the isotropic case from Andersen.

3.1.9 Replica exchange with solute tempering

Replica exchange molecular dynamics method (Usually it is called REM) [267, 268, 269] has been widely used to accelerate conformational sampling of complex systems, such as biomolecules. However, the number of replicas increases with the square root of the number of degrees of freedom [268, 270], which limits its applications for large systems. To overcome this problem, Liu *et al.* [176] proposed a so-called "replica exchange with solute tempering" (REST) variant of REM, which is able to reduce significantly the number of replicas.

This scheme is based on the observation that the simulation box of most biomolecular systems consists of relatively small solute molecules (proteins, DNA...) and a large number of solvent molecules. However, the main interest is usually in the solute conformations. Therefore, REST aims at enhancing only the dynamics of the solute molecules.

In a standard REM simulation, multiple replicas of the same system are prepared. Each replica has different temperature and/or Hamiltonian and is sampled independently with the canonical ensemble MD simulations for some time. Then, with a certain frequency, exchange between neighboring replicas is attempted with some probabilities. These exchange probabilities are determined so that the ensemble of each replica obeys the Boltzmann distribution.

In case the replicas differ only in the Hamiltonian, the Boltzmann distribution for each replica is given by

$$P_i(X_i) = Z_i^{-1}\exp\left[-\beta E_i(X_i)\right], \tag{3.132}$$

where E_i is the energy and Z_i is the partition function of replica i. The exchange probability $W\left(i \to j\right)$ between replica i and j is determined from the detailed balance condition

$$P_i(X_i)P_j(X_j)W(i \to j) = P_i(X_j)P_j(X_i)W(j \to i), \tag{3.133}$$

which gives the ratio of the transition probabilities

$$\frac{W(i \to j)}{W(j \to i)} = \exp(-\Delta), \tag{3.134}$$

with

$$\Delta = \beta\left[E_i(X_j) + E_j(X_i) - E_i(X_i) - E_j(X_j)\right]. \tag{3.135}$$

If the Metropolis criterion is applied, the acceptance probability is chosen as

$$W(i \to j) = \begin{cases} 1 & \text{if} \Delta \leq 0 \\ \exp(-\Delta) & \text{if} \Delta > 0. \end{cases} \tag{3.136}$$

In the REST variant proposed by Liu *et al.* [176], the system of interest (e.g. protein solvated in water) is therefore divided into two parts: the protein molecule (denoted by p) and water molecules (denoted by w). The potential energy of each replica used to enhance sampling is scaled such that the protein molecule in different replicas has different effective temperature while the solvent's effective temperature is the same in all replicas. The potential energy can be divided into three terms

$$V = V_{pp} + V_{pw} + V_{ww}, \tag{3.137}$$

where the first term represents the energy of the protein, the second one the interaction energy between the protein and water, and the third, the interaction of water molecules with each other. The system of interest is usually chosen to be the lowest replica, i.e. the protein solution at the inverse temperature $\beta_0 = \frac{1}{k_B T_0}$ with the energy potential given by Eq. (3.137). The energy potential for the highest temperature replica is given by

$$V_H = \frac{\beta_H}{\beta_0} V_{pp} + \sqrt{\frac{\beta_H}{\beta_0}} V_{pw} + V_{ww}, \tag{3.138}$$

where β_H is the inverse of the highest effective protein temperature. The energy potentials for the intermediate replicas are defined as the linear combination of the two-end potentials defined above

$$\begin{aligned} V_i &= (1 - \lambda_i) V + \lambda_i V_H \\ &= \left[(1 - \lambda_i) + \frac{\beta_H}{\beta_0} \lambda_i \right] V_{pp} + \left[(1 - \lambda_i) + \sqrt{\frac{\beta_H}{\beta_0}} \lambda_i \right] V_{pw} + V_{ww}, \end{aligned} \tag{3.139}$$

where λ_i is related to the effective inverse temperature β_i of replica i through the formula

$$\lambda_i = \frac{\beta_i - \beta_0}{\beta_H - \beta_0}. \tag{3.140}$$

The energy potentials of all the replicas are now completely defined by equations (3.137)-(3.139). One can directly use the Hamiltonian replica exchange routine of GROMACS program [203], which is based on the λ–dynamics to perform REST simulations. The number of replicas and values of effective temperature can be estimated using the approach of ref. [271], in which only the number of protein's atoms is taken into account. The relation (3.140) is then used to determine the values of the coupling constants λ_i.

3.2 Computational spectroscopy

In this thesis, extensive comparison with experimental spectroscopic properties is made. Here, we present the theoretical framework of the spectroscopy calculations carried out in this work.

3.2.1 NMR chemical shifts

In NMR experiments of a solute in aqueous solution, the sample is exposed to an external magnetic field $\vec{B}^{\text{ext}}$. This magnetic field induces an electronic current $\vec{j}(\vec{r})$. It, in turn, produces an additional inhomogeneous magnetic field $\vec{B}^{\text{ind}}$ at the positions of sample's nuclei. The latter field is sensitive to the electronic density and thus different for different nuclei depending on their chemical environment. Therefore, it determines the resonance frequencies of the nuclear spins measured in an NMR experiment [273] by irradiating the sample with an electromagnetic pulse.

The chemical shielding tensor $\overleftrightarrow{\sigma}$ is defined as the proportionality tensor factor between the induced and the external magnetic field at the positions of the nuclei

$$\overleftrightarrow{\sigma}(\vec{r}) = \frac{\partial \vec{B}^{\text{ind}}(\vec{r})}{\partial \vec{B}^{\text{ext}}}. \tag{3.141}$$

NMR experiments of biomolecules in solution measure isotropic shieldings. This can be calculated as the trace of the shielding tensor $\overleftrightarrow{\sigma}$

$$\sigma = \frac{1}{3}\text{Tr}\left(\overleftrightarrow{\sigma}\right). \tag{3.142}$$

The shielding tensor is sensitive to the local electronic density and thus to the chemical bonding environment. The resonance condition is given by

$$\Delta E = h\nu = g_H \mu_H (1-\sigma)\vec{B}^{\text{ext}}. \tag{3.143}$$

ΔE is the resonance energy; ν is the corresponding frequency, h is Planck's constant. g_H and μ_H are the nuclear g-factor and the nuclear magneton, respectively. The chemical shift δ of a given sample is defined with respect to a reference frequency

$$\delta = \frac{v_{\text{sample}} - v_{\text{reference}}}{v_{\text{reference}}} \times 10^6\,[\text{ppm}], \tag{3.144}$$

where ν_{sample} is the frequency of a nucleus in the sample, and $\nu_{\text{reference}}$ is that of the reference system.

From this discussion it is clear that the key quantity for NMR chemical shift calculations is the local magnetic field $\vec{B}^{\text{ind}}$ that is induced by the total electronic current $\vec{j}(\vec{r})$. It reads

$$\vec{B}^{\text{ind}}(\vec{r}) = \frac{\mu_0}{4\pi}\int d^3\vec{r}' \frac{\vec{r}' - \vec{r}}{\left|\vec{r}' - \vec{r}\right|^2} \times \vec{j}(\vec{r}'), \tag{3.145}$$

where μ_0 is the permeability in vacuum. The total electronic current $\vec{j}(\vec{r})$ can be calculated using first principles density functional perturbation theory [208]. The Kohn-Sham orbitals are expanded in powers of the external magnetic field

$$\phi_i = \phi_i^{(0)} + B^{ext}\phi_i^{(1)} + \left(B^{ext}\right)^2\phi_i^{(2)} + ... \tag{3.146}$$

They minimize the total density functional

$$E_{KS}^{\text{tot}} = E_{KS}\left[\phi_i^{(0)} + B^{ext}\phi_i^{(1)} + ...\right] + \varepsilon^{\text{pert}}\left[\phi_i^{(0)} + B^{ext}\phi_i^{(1)} + ...\right], \tag{3.147}$$

where E_{KS} is the unperturbed Kohn-Sham density functional in Eq. (3.50). The corresponding unperturbed effective Kohn-Sham Hamiltonian is

$$H_{KS} = -\frac{\hbar^2}{2m_e}\nabla^2 + v_{\mathrm{ks}}(\rho), \tag{3.148}$$

where v_{ks} is the Kohn-Sham potential given by Eq. (3.52).

The external magnetic field can be written as

$$\vec{B}^{\mathrm{ext}} = \vec{\nabla} \times \vec{A}\,(\vec{r}\,), \tag{3.149}$$

where $\vec{A}(\vec{r})$ is the vector potential. The vector potential is not uniquely defined since the gradient of any scalar function can be added without changing $\vec{B}^{\mathrm{ext}}$. In the case of a homogeneous magnetic field (as in the case of NMR experiments), $\vec{A}$ can be conveniently chosen as [273]

$$\vec{A}\,(\vec{r}\,) = -\frac{1}{2}\left(\vec{r} - \vec{R}^{\mathrm{go}}\right) \times \vec{B}^{\mathrm{ext}}, \tag{3.150}$$

where $\vec{R}^{\mathrm{go}}$ is the gauge origin of the vector potential $\vec{A}(\vec{r})$, *i.e.* the center of the vector potential. This is an arbitrary parameter.

The main point in defining $\vec{A}(\vec{r})$ is that the introduction of the external magnetic field $\vec{B}^{\mathrm{ext}}$ is equivalent to replacing the momentum operator $\vec{p} = -i\hbar\vec{\nabla}$ in the unperturbed Kohn-Sham Hamiltonian (3.148) by $\vec{p} - \frac{e}{c}\vec{A}\,(\vec{r}\,)$, where c is the speed of light. This leads to a perturbation term in the Hamiltonian at the first order in the field strength (paramagnetic term), $-\frac{e}{m_e c}\vec{p}\cdot\vec{A}\,(\vec{r}\,)$ and one at the second order (diamagnetic term), $\frac{e^2}{2m_e c^2}\vec{A}\,(\vec{r}\,)\cdot\vec{A}\,(\vec{r}\,)$:

$$H_{KS}^{\mathrm{tot}} = H_{KS} + H^{\mathrm{pert}}, \tag{3.151}$$

with

$$H^{\mathrm{pert}} = -\frac{e}{m_e c}\vec{p}\cdot\vec{A}\,(\vec{r}\,) + \frac{e^2}{2m_e c^2}\vec{A}\,(\vec{r}\,)\cdot\vec{A}\,(\vec{r}\,). \tag{3.152}$$

Then, the perturbation functional energy $\varepsilon^{\mathrm{pert}}$ reads

$$\begin{aligned}\varepsilon^{\mathrm{pert}} = \sum_i \left\langle \phi_i \left| H^{\mathrm{pert}} \right| \phi_i \right\rangle = \sum_i \frac{e^2}{2m_e c^2}\left\langle \phi_i^{(0)} \left| \vec{A}\,(\vec{r}\,)\cdot\vec{A}\,(\vec{r}\,) \right| \phi_i^{(0)} \right\rangle \\ - \frac{e}{m_e c}\left(\left\langle \phi_i^{(1)} \left| \vec{p}\cdot\vec{A}\,(\vec{r}\,) \right| \phi_i^{(0)} \right\rangle + \left\langle \phi_i^{(0)} \left| \vec{p}\cdot\vec{A}\,(\vec{r}\,) \right| \phi_i^{(1)} \right\rangle\right),\end{aligned} \tag{3.153}$$

in which $\phi_i^{(0)}$ are the ground state Kohn-Sham orbitals and $\phi_i^{(1)}$ are the first order corrections.

The first order perturbation gives rise to a correction $\phi^{(1)}$ to the electronic ground state in Eq. (3.146). This correction is responsible for the induced current, which is obtained from the functional Eq. (3.153) as a sum of a dia- and a paramagnetic term [208, 209]:

$$\begin{aligned}\vec{j}\,(\vec{r}\,) = \frac{e^2}{m_e}\vec{A}\,(\vec{r}\,)\sum_i \left|\phi_i^{(0)}\,(\vec{r}\,)\right|^2 \\ - \frac{e}{m_e}\sum_i \left[\left\langle \phi_i^{(1)} \left| \vec{p} \right| \phi_i^{(0)} \right\rangle - \left\langle \phi_i^{(0)} \left| \vec{p} \right| \phi_i^{(1)} \right\rangle\right].\end{aligned} \tag{3.154}$$

Both contributions depend individually on the choice of the gauge. To address the implications resulting from the ambiguity in such choice, several approaches have been developed, see e.g. [273]. Here, we use the so-called "$\vec{R}^{\mathrm{go}} = \vec{r}$" variant of the "Continuous Set of Gauge Transformations" (CSGT) method [274], which is most appropriate for plane-wave based calculations. For each point in space $\vec{r}$, the current density (Eq. 3.154) is calculated by setting $\vec{R}^{\mathrm{go}}$ equal to $\vec{r}$.

For heavy atoms such as Pt, it is essential to take into account relativistic effects because the core electrons (close to the nucleus) in these atoms move at a considerable fraction of the speed of light. In this thesis, we have used the Zero Order Regular Appoximation (ZORA) appoach to include the relativistic correction in the calculations of Pt NMR shielding constants. In the ZORA approach, the effective one-electron Kohn-Sham Hamiltonian (3.148) is replaced by the relativistic Hamiltonian for an electron:

$$H_{\mathrm{KS}}^{\mathrm{ZORA}} = v_{\mathrm{ks}} + \vec{p}\frac{c^2}{2c^2 - v_{\mathrm{ks}}}\vec{p} + \frac{c^2}{2c^2 - v_{\mathrm{ks}}}\overleftrightarrow{\sigma}\left(\vec{\nabla}v_{\mathrm{ks}} \times \vec{p}\right), \tag{3.155}$$

where v_{KS} is the effective one-electron Kohn-Sham potential given in Eq. (3.52), $\overleftrightarrow{\sigma}$ is the Pauli matrix, c is the speed of light. The magnetic field can be then included in the same way, *i.e.*, by means of the substitution

$$\vec{p} \to \vec{p} - \frac{e}{c}\vec{A}(\vec{r}). \tag{3.156}$$

3.2.2 Extended X-ray absorption fine structure (EXAFS)

EXAFS spectroscopy probes the local environment of metal ions in biological systems [275, 220]. It provides detailed local structural information such as interatomic distances and coordination numbers. In an EXAFS experiment, an absorbed X-ray photon liberates an electron from a core level resulting in an absorption edge in the spectrum. The energies of the edges are unique to the type of atom. Well above the absorption edge ($\geq$ 30 ev), oscillatory wiggles arise from higher order multiple scattering. In fact, the ejected photoelectron wave function scatters from the atoms surrounding the absorbing atom, thereby creating interferences between the outgoing and scattered waves. Thus, the signal is sensitive to the short-range order in the system under investigation. These quantum interferences cause an energy-dependent variation of the X-ray absorption coefficient μ, a measurable quantity. The EXAFS function χ of a specific atom is then defined as the normalized, oscillatory part of the X-ray absorption above the absorption edge of that atom

$$\chi(E) = \frac{\mu(E) - \mu_0(E)}{\mu_0(E)}, \tag{3.157}$$

where E is the energy of the photoelectron, defined with reference to a core level E_i from which the electron is liberated

$$E = \hbar\omega + E_i. \tag{3.158}$$

$\hbar\omega$ is the energy of X-ray photon. $\mu_0(E)$ is the absorption coefficient of the isolated atom. Eq. (3.157) is the theoretical definition for the EXAFS, in which $\mu_0(E)$ can be

calculated theoretically [276] but it is not accessible experimentally. In experiments, the EXAFS is instead calculated from data reduction using the expression [275]

$$\chi(E) = \frac{\mu_{\exp}(E) - \mu_b(E)}{\Delta\mu_{\exp}}, \tag{3.159}$$

where $\mu_{\exp}(E)$ is the experimentally measured absorption, $\Delta\mu_{\exp}$ an estimate of the edge step, and $\mu_b(E)$ is a smooth background curve that is fit to the experimental data.

$\mu(E)$ in Eq. (3.157) is proportional to the transition rate as given by Fermi's golden rule within the dipole approximation [220]

$$\mu(E) \propto \sum_f |\langle\psi_f|\hat{\varepsilon}\cdot\vec{r}|\psi_i\rangle|^2 \delta(E - E_f), \tag{3.160}$$

where $|\psi_i\rangle$ is the initial deep core state with energy E_i and $|\psi_f\rangle$ is the final state with energy E_f, which is initially unoccupied. $\hat{\varepsilon}$ is the photon's electric polarization vector. In this work, the calculations of EXAFS spectra have been done using multiple scattering approach [219, 225, 220]. In the Green's function formulation, $\mu(E)$ can be rewritten in terms of the imaginary part of the one-particle Green's function $G(\vec{r}, \vec{r}', E)$ describing all of the possible ways for a photoelectron to interact with the surrounding atoms

$$\mu(E) \propto \frac{-1}{\pi}\mathrm{Im}\,\langle\psi_i|\hat{\varepsilon}^*\cdot\vec{r}G(\vec{r},\vec{r}',E)\,\hat{\varepsilon}\cdot\vec{r}'|\psi_i\rangle\,\Theta(E - E_F), \tag{3.161}$$

where $\Theta(E - E_F)$ is the broadened step function at the Fermi's energy level E_F.

The Green's function can be expanded in a multiple scattering series in which each term contains a successive product of free propagators between geometric sites and scattering matrices at each site. The scattering potential seen by the photoelectron is approximated by the muffin-tin potential in which the atomic potentials are approximated as spherically symmetric out to a finite radius, and the interstitial potential between the atomic sphere is approximated as a constant.

Instead of representing χ as a function of energy, the wave vector, $k = \sqrt{2m_e E/\hbar^2}$ is used. $\chi(k)$ is then expressed as a function of the contributions from each of the scattering paths [225, 220]

$$\begin{aligned}\chi(k) &= \sum_\Gamma \chi_\Gamma(k) \\ &= \sum_\Gamma \frac{S_0^2 |F_\Gamma(k)|}{kR_\Gamma^2}\sin(2kR_\Gamma + \Phi_\Gamma(k) + 2\delta_c(k))\, e^{-2\sigma_\Gamma^2 k^2} e^{-2R_\Gamma/\lambda(k)},\end{aligned} \tag{3.162}$$

where Γ runs over all possible multiple scattering paths Γ within a cluster of atoms. $|F_\Gamma(k)|$ is the effective scattering amplitude for path Γ [220]. S_0^2 is an amplitude reduction factor to account for intrinsic losses and interference. R_Γ is the length of the path Γ. $\Phi_\Gamma(k)$ is the net scattering phase shift, and $\delta_c(k)$ is the final-state phase shift at the absorbing atom. $\sigma_\Gamma^2 = \langle(\delta R_\Gamma)\rangle$ is the mean square variation in R_Γ, and $e^{-2\sigma_\Gamma^2 k^2}$ is the path Debye-Waller factor, which accounts for the fluctuations of the path length due to structural and thermal disorder. $\lambda(k)$ is the mean free

path of the photoelectron and hence, $e^{-R_\Gamma/\lambda(k)}$ accounts for the finite life-time of the photoelectron.

The computation of EXAFS spectra using Eq. (3.162) has been implemented in the FEFF code developed by Rehr *et al.* [218, 219]. First, we calculate the electronic structure of the isolated atoms by solving the Dirac-Fock equation [Za02, Bu10]. The atomic charge densities are then superimposed in the desired geometry, and averaged to create the muffin-tin potential. Hedin-Lundqvist self-energy terms are added to account for inelastic losses. Then, $\lambda(k)$ is calculated from the imaginary part of the photoelectron self-energy and the core-hole lifetime [220]. $\delta_c(k)$ are calculated by integration within the atomic potential outward to the radius, and matching the boundary conditions to free spherical waves between the atoms [220].

To calculate the set of the scattering amplitudes $F_\Gamma(k)$ as a function of k, a multiple scattering expansion is done. FEFF uses the scattering matrix formalism of Rehr and Albers [220] in which the effective scattering amplitude for each path is given by

$$F_\Gamma = \mathrm{Tr} M \cdot F^N \cdots F^2 F^1, \tag{3.163}$$

where M is the termination matrix for the final state at the absorption site. F^i is the scattering matrix at site i. A huge number of paths enter into the calculation, but many of them are (nearly) equivalent to each other and their calculation can be omitted. A path filter is implemented in FEFF to estimate the importance of each path, defined as the integral over the full k range of $|\chi_\Gamma(k)|\,dk$. Any path with a value of the integral larger than a few percent of the largest path will be used in the calculation of $\chi(k)$.

To calculate the Debye-Waller factor as defined above, FEFF uses the correlated Debye model [277], in which the mean-square displacement is

$$\sigma_\Gamma^2(T) = \frac{\hbar}{M} \int_0^{\omega_{\max}} \frac{d\omega}{\omega} \rho_\Gamma(\omega) \coth\left(\frac{\beta\hbar\omega}{2}\right), \tag{3.164}$$

where T is the temperature, M is the atomic mass, ω is the vibrational frequency, $\beta = 1/k_B T$, and ρ_Γ is the projected density of vibrational states. In this model, ρ_Γ is given by [220]

$$\rho_\Gamma(\omega) = \frac{3\omega^2}{\omega_D^2}\left[1 - \frac{\sin(\omega R_\Gamma/c)}{\omega R_\Gamma/c}\right]. \tag{3.165}$$

$\omega_D = k_B\theta_D/\hbar$ is the Debye frequency, θ_D is the Debye temperature, $c = \omega_D/k_D$ is the Debye approximation for the speed of sound, $k_D = 6\pi^2 N/V$, and N/V is the atomic density number.

3.2.3 Circular dichroism (CD)

Circular dichroism (CD) spectra can be used to analyze the secondary structure of proteins. As all chiral groups, protein backbone units absorb left- and right-circularly polarized light in a different manner. Hence, the extinction coefficients of the absorption of left- (ε_L) and right (ε_R)-circularly polarized light differ: $\varepsilon_L \neq \varepsilon_R$.

The CD function $\Delta\varepsilon(\lambda)$ is the difference in the absorption of left- and right-circularly polarized light, plotted as a function of the wavelength λ of the polarized light

$$\Delta\varepsilon\left(\lambda\right) = \varepsilon_L\left(\lambda\right) - \varepsilon_R\left(\lambda\right). \tag{3.166}$$

The integral of $\Delta\varepsilon(\lambda)$ over a wavelength range associated with a particular transition from the ground state 0 to an excited state k is the rotational strength of that transition (R^{0k}). The latter can be calculated using the Rosenfeld equation [278]

$$R^{0k} = \mathrm{Im}\left(\left\langle\psi^0\left|\vec{\mu}\right|\psi^k\right\rangle\left\langle\psi^k\left|\vec{m}\right|\psi^0\right\rangle\right), \tag{3.167}$$

where, $\vec{\mu}$ is the electric, and $\vec{m}$ the magnetic transition dipole moment operator. ψ^0 and ψ^k are the wave functions of the ground state and the excited state, respectively.

For proteins, which have a very large number of atoms and electrons, approximations are needed to calculate ψ^0 and ψ^k. In the matrix method [279], the protein is divided into M chromophores with a certain number of local excitations at each chromophore. The approach is based on coupled-oscillator and exciton models, where the Hamiltonian is described by energies and interactions between the transitions on the chromophores. For protein CD spectroscopy, the most important chromophores are the amide groups. The relevant excitations are $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions of the amide carbonyl groups.

In the matrix method, the total wave function of the protein is expressed as a linear superposition of basis functions Φ_{ia}. Each basis function is a product of M monomer wave functions. Electronic excitations may occur within a group but not between groups. The basis function is further restricted to allow only one group to be excited [206, 280, 281, 207]. Thus:

$$\Phi_{ia} = \varphi_{10} \cdot \cdots \cdot \varphi_{ia} \cdot \cdots \cdot \varphi_{M0}, \tag{3.168}$$

where φ_{ia} represents the wave function of chromophores i, which has undergone an electronic excitation $\varphi_{i0} \rightarrow \varphi_{ia}$. The excited state wave function of the protein is then given by

$$\psi^k = \sum_i^M \sum_a^{n_i} c_{ia}^k \Phi_{ia}. \tag{3.169}$$

The sum above is over M chromophores and n_i excitations on each chromophore. c_{ia}^k's are expansion coefficients to be determined by solving the Schrödinger equation for the excited state k of the protein [280, 281]

$$H\psi^k = E^k\psi^k. \tag{3.170}$$

The total Hamiltonian H for the protein consisting of M chromophores reads

$$H = \sum_{i=1}^{M} H_i + \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} V_{ij}, \tag{3.171}$$

where H_i is the Hamiltonian of chromophore i and V_{ij} is the interaction potential between chromophores.

The Hamiltonian (3.171) can be written in matrix form, in which the diagonal elements are the transition energies for each transition of each chromophore, and

the off-diagonal elements are the interaction energies between different transitions. These interactions are the cause of the dependency of protein CD on secondary structure.

The diagonal elements of the Hamiltonian matrix are calculated as

$$H_{ia} = E_{ia} = \langle \Phi_{ia} | \, H_i \, | \Phi_{ia} \rangle \, . \tag{3.172}$$

The off-diagonal elements of the Hamiltonian matrix are calculated as

$$V_{ia,jb} = \langle \Phi_{ia} | \, V_{ij} \, | \Phi_{jb} \rangle \, , \tag{3.173}$$

e.g. for the interaction between the transition $0 \rightarrow a$ on chromophore i and the transition $0 \rightarrow b$ on chromophore j, it reads

$$V_{i0\rightarrow a;j0\rightarrow b} = \int \int \varphi_{i0} \varphi_{ia} V_{ij} \varphi_{j0} \varphi_{jb} d\tau_i d\tau_j . \tag{3.174}$$

This interaction can be regarded as an electrostatic interaction between charge densities ρ. In this case Eq. (3.174) becomes

$$V_{i0\rightarrow a;j0\rightarrow b} = \int_{r_i} \int_{r_j} \frac{\rho_{i0a}\left(r_i\right) \rho_{j0b}\left(r_j\right)}{4\pi\varepsilon_0 r_{ij}} d\tau_i d\tau_j , \tag{3.175}$$

where $\rho_{i0a}\left(r_i\right)$ and $\rho_{j0b}\left(r_j\right)$ denote the transition electron densities on chromophores i and j, respectively, ε_0 is the vacuum permittivity and r_{ij} is the distance between the chromophores.

In the "monopole-monopole" approximation, used here, the charge and transition densities are approximated by point charges and the matrix element $V_{i0\rightarrow a;j0\rightarrow b}$ is re-cast as a sum of Coulomb interactions

$$V_{i0\rightarrow a;j0\rightarrow b} = \sum_{s=1}^{N_s} \sum_{t=1}^{N_t} \frac{q_s q_t}{r_{st}} , \tag{3.176}$$

where q_s and q_t are point charges on chromophores i and j. N_s and N_t are the number of these charges on the chromophores. Thus the magnitude and the location of these monopoles reflect the magnitude and the orientation of the transition moments. They are critical for determining the inter-chromophore interactions.

The computation of CD spectra requires atomic coordinates and the position of the chromophores as input. Each chromophore is parameterized by a set of monopoles, which are superposed on the chromophoric groups in the protein, and describe their electrostatic potential. There are commonly used parameter sets for the matrix method: NMAabinit [207] that has been derived entirely from *ab initio* calculations and NMAsemi [282] that has been derived from a combination of semiempirical calculations and experimental data. By calculating the interaction between all the different electronic excitations, the Hamiltonian matrix is constructed. Diagonalization of the Hamiltonian matrix yields the energies (the eigenvalues) for all states k and the corresponding wave functions ψ^k.

Rotational strengths of the transitions R^{0k}, *i.e.* their spectral intensities, are calculated by using Eq. (3.167). This yields a line spectrum. In an experimental

spectrum the transitions are broadened by various mechanisms, *e.g.* by lifetime broadening, vibronic coupling or interactions with the environment. Therefore, the CD spectrum can be constructed from the rotational strengths by convoluting them with normalized band shape functions such as a Gaussian or Lorentzian line shape function. For Gaussian band shape functions, which typically provide better results than Lorentzian functions, the CD spectrum reads [280]

$$\Delta\varepsilon\left(\lambda\right) \propto \sum_{k} \frac{R^{0k}}{\Gamma_k} \exp\left[-\left(\frac{\lambda-\lambda_k}{\Gamma_k}\right)^2\right], \tag{3.177}$$

where Γ_k is the half-width of the Gaussian.

4

Structural determinants of cisplatin and transplatin binding to Ctr1's Met-rich motif: a computational spectroscopy approach

This chapter is reproduced from Ref. [63]

Abstract

The cellular uptake of cisplatin and of other platinum-based drugs is mediated by the high-affinity copper transporter *Ctr1*. The 8-residue long peptide called *Mets7* (which has the amino acid sequence MTGMKGMS) mimics one of *Ctr1*'s extracellular methionine(Met)-rich motifs. It is an excellent model for investigating the interaction of platinum drugs with *Ctr1* in *in vitro* and *in vivo* conditions. Arnesano *et al.* have shown that: (i) Cisplatin loses all of its ligands upon reaction with *Mets7* and the metal ion binds to the three Met residues and completes its coordination shell with a fourth ligand which can be a chloride or a water/hydroxyl oxygen. (ii) Transplatin loses only the chlorido ligands, which are replaced by Met residues. Here, we provide information on the structural determinants of cisplatin-*Mets7* and transplatin-*Mets7* adducts by computational methods. The predictions are validated against EXAFS, NMR, and CD spectra. While EXAFS gives information restricted to the metal coordination shell, NMR provides information extended to residue atoms around the coordination shell and, finally, CD provides information about the overall conformation of the peptide. This allows us to elucidate the different reaction modes of cisplatin and transplatin towards the peptide, as well as to propose the platinated peptides $[Pt^{2+}X]$-(M*TGM*KGM*S) (X = Cl^-, OH^-) and trans$[Pt(NH_3)_2]^{2+}$-(M*TGM*KGMS) as the most relevant species occurring in water solution.

4.1 Introduction

Cisplatin [cis-diamminedichloridoplatinum(II)] is one of the most widely used anticancer drugs. It is active against many types of cancer, such as testicular, ovarian, cervical, and colorectal cancers and relapsed lymphoma [87, 2]. Unfortunately, the drug suffers from several drawbacks, such as side effects and intrinsic or acquired resistance, which severely limit the efficacy of the drug [87, 3, 1]. The cellular mechanisms behind the resistance of cisplatin and other platinum-based drugs are multifactorial and not completely understood [30, 2, 3]. The generally accepted mechanism involves an increased cisplatin detoxification, an improved repair of the DNA lesions, and a decreased accumulation of the drug [30, 28]. Among these factors, the reduced drug accumulation is a common and distinctive feature of resistant cells [30, 131, 2]. It is, therefore, important to characterize the molecular mechanisms of cisplatin's efflux and cellular uptake. The first involves the copper pumps *ATP7A* and *ATP7B* in the secretory vesicles [153, 30, 132] and the cytosolic copper chaperone *Atox1* [110, 34]. All of these proteins can bind cisplatin and enhance its efflux/sequestration [30, 132, 110, 34]. Cisplatin may be also ensnared in melanosomes that are subsequently exported from the cells [177]. The cellular uptake is governed, at least in part, by the homotrimeric copper transporter *Ctr1* [30, 178, 110]. *Ctr1* forms a channel-like pore in the membrane [179], featuring intra- and extracellular domains. Flow of the active platinum substrate through the channel-like pore of the *Ctr1* trimer has been proposed [180, 181]. Alternatively, cisplatin could form coordination bonds with the protein [142, 143]. The extracellular N-terminal domain contains methionine(Met)-rich motifs, which are able to bind copper [182]. Met's sulfur is indeed a good donor atom for platinum [183]. Hence, some cisplatin could bind to the extracelular Met-rich motifs of *Ctr1* and an endocytic vesicle might be formed. In this way the vescicle could incorporate cisplatin present in the extracellular solution while protecting the drug from deactivation and transporting it to subcellular compartments [142, 143].

To provide insights on *Ctr1*–cisplatin interactions, Arnesano *et al.* have carried out *in vitro* studies on reactions of cisplatin, the inactive isomer transplatin [transdiamminedichloridoplatinum(II)], and other platinum complexes with the MTGMKGMS peptide (*Mets7*) [142, 143]. The peptide mimics the seventh Met-rich motif of yeast *Ctr1*. Circular dichroism (CD) measurements showed that *Mets7* features a β-turn conformation after reacting with cisplatin, whilst it is a random coil in the Apo form [142]. Furthermore, electrospray mass spectrometry (ESI-MS) data showed that cisplatin progressively loses all ammine and chlorido ligands and ultimately coordinates only to *Mets7*. This contrasts with transplatin, which keeps the two ammine ligands. Consistently, ^{1}H and ^{13}C nuclei of ϵ-CH_3 and γ-CH_2 of methionines feature a large downfield shift in NMR spectra upon platinum coordination [142]. Other groups also reported a similar binding mode of cisplatin to partial or full-length N-terminal domain of human *Ctr1* [146, 147]. Hence, *Mets7* is a good mimic of the whole N-terminal domain of *Ctr1* in terms of reactivity against cisplatin.

The structural basis of cisplatin and transplatin binding to *Mets7* is still lacking. Herein, we have addressed this issue by using computational tools. Computational approaches are indeed widely used to investigate cisplatin and its interactions with biomolecules [45]–[62], [155]–[165]. Specifically, we have used hybrid Car-Parrinello density functional theory-based QM/MM simulations [166, 167, 168, 169, 170, 171,

172] on models of platinated *Mets7*, extracted from available experimental information [142, 143]. These models contain the platinum moieties ($[PtCl]^+$, $[Pt(H_2O)]^{2+}$, $[Pt(OH)]^+$ and *trans*-$[Pt(NH_3)_2]^{2+}$). The QM/MM simulations, albeit rather accurate, can cover only rather short time scales (0.1 ns or less). Whilst these are likely to be enough to provide valuable information on the local platinum coordination, global conformational properties of the peptide are out of reach. To extend the time-scale of the simulations, we have developed AMBER-type [173] force field parameters of the platinum moieties based on the QM/MM simulations via the so-called "force matching" procedure [174]. The parameters are derived so as to reproduce the QM/MM forces, the electrostatic potential, and the electric field obtained by QM/MM simulations. This procedure was already applied successfully to platinated DNA [175]. Finally, 0.2μs-long replica exchange simulations [176] based on these force field parameters are performed for the platinated peptides. The accuracy of the predicted models is established by a comparison with the NMR and CD data already reported in the literature [142, 143], as well as EXAFS spectra measured expressly for this study.

4.2 Computational protocol

4.2.1 Molecular models

We have considered in total seven models: one free *Mets7* peptide (called Apo hereafter), three cisplatn-*Mets7* (C1–C3) and three transplatin-*Mets7* (T1–T3) adducts (Fig. 4.1).

The cisplatn-*Mets7* adducts include $[Pt^{2+}X]$-(M*TGM*KGM*S), in which X = Cl^- (C1); H_2O (C2); OH^- (C3).

The transplatin-*Mets7* adducts are *trans*-$[Pt(NH_3)_2]^{2+}$-(M*TGM*KGMS) (T1); *trans*-$[Pt(NH_3)_2]^{2+}$-(M*TGMKGM*S) (T2); *trans*-$[Pt(NH_3)_2]^{2+}$-(MTGM*KGM*S) (T3) ("*" indicates the donated methionine).

Apo was built as a linear chain by using the Leap package [185].

In the cisplatin-*Mets7* adducts (C1–C3, Fig. 4.1), the platinum coordination was taken from DFT-based optimized model compounds (see Appendix A.1). The platinum coordination's donor atoms are the *Mets7*'s three methionine sulfurs and either a chloride (C1), or a water oxygen (C2), or a hydroxyl oxygen (C3).

In the transplatin-*Mets7* adducts (T1–T3, Fig. 4.1), the donor atoms are two ammine nitrogens and two sulfur atoms of either Met1 and Met4 (T1), or Met1 and Met7 (T2), or Met4 and Met7 (T3). The six systems, along with Apo, were inserted in boxes containing water molecules and chloride counterions (Tab 4.1).

The platinated (C1–C3, T1–T3) and Apo peptides were subjected to 0.4 μs of preparatory AMBER-based molecular dynamics (MD) simulations using the NAMD program [186]. The AMBER parm99 force field [173] with the "Stony Brook" modification for the backbone torsions [190], the TIP3P water model [192] and the Smith&Dang force field [193] were used for the peptide, water molecules, and counterions, respectively. The bond distances and angles were constrained using the

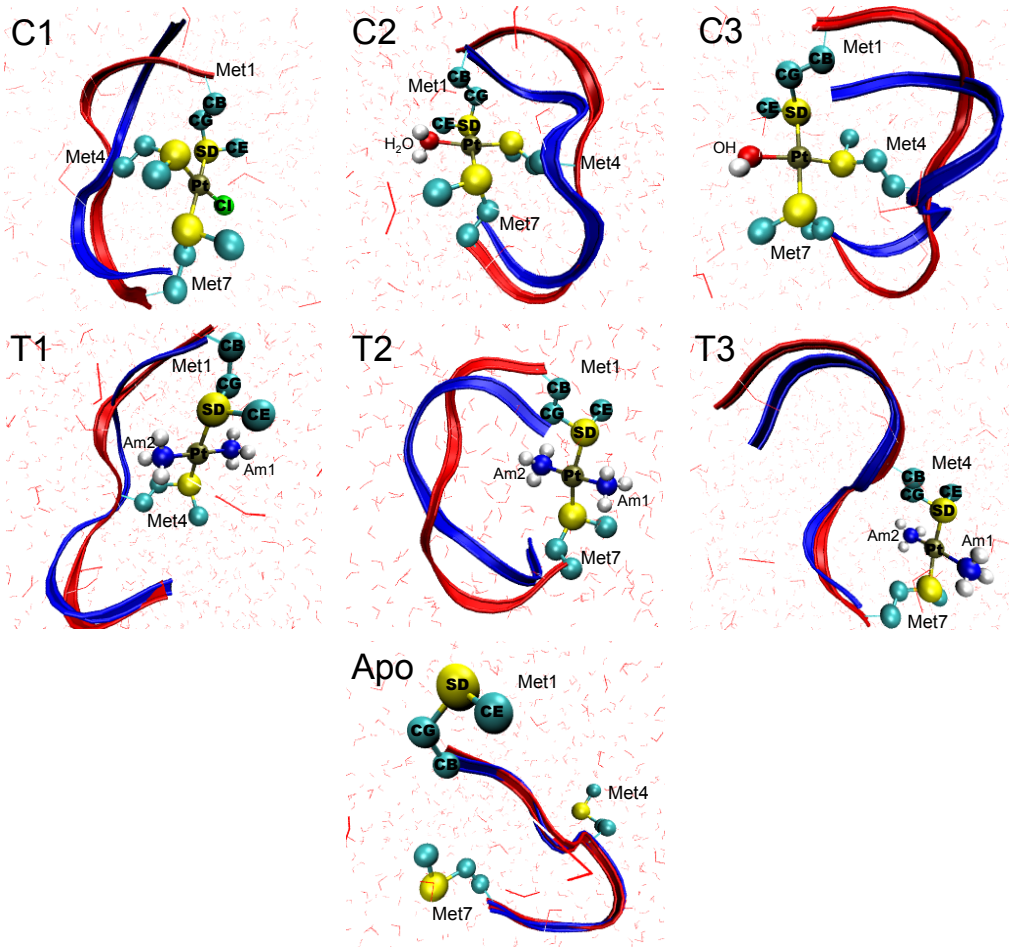


Figure 4.1 Structural models of platinated and Apo peptides emerging from REST and QM/MM simulations. The backbone conformations of the last QM/MM snapshot (in red ribbons) and a representative REST snapshot (in blue ribbons) are superimposed. The QM atoms are shown in stick-and-balls (CB=C_β, CG=C_γ, CE=C_ϵ). Water is shown in lines. Counterions are not shown for the sake of clarity.

SHAKE algorithm [235] to the values taken from $[Pt(SH_2)_3Cl]^+$, $[Pt(SH_2)_3(H_2O)]^{2+}$, $[Pt(SH_2)_3(OH)]^+$ and *trans*-$[Pt(SH_2)_2(NH_3)_2]^{2+}$ gas-phase-optimized structures (Tab. A.1 in Appendix). Long-range electrostatic interactions were evaluated using the Particle Mesh Ewald (PME) method [236] The cutoff for the real part of the PME and for the van der Waals interaction was set to 10 Å. An integration time step of 2fs was used. Constant temperature (300 K) and pressure (1 atm) conditions were achieved by using Langevin dynamics with a coupling coefficient of 5 ps and the Nose-Hoover Langevin piston algorithm [237] with an oscillation period of 200 fs and the damping timescale of 100 fs. A cluster analysis [187] allowed us to identify representative structures of the most populated clusters. It turned out that only one cluster was sufficient to comprise 70% or more of the conformational space explored by the MD trajectories. Hence, we considered only the representatives of the most populated clusters.

	Apo	C1	C2	C3	T1	T2	T3
No. atoms	7760	9488	8783	8451	10644	9924	10707
No. waters	2547	3123	2886	2676	3505	3265	3526
No. counter ions	1	2	3	2	3	3	3
QM-part	Met1, 2, 3	$[PtCl]^+$, Met1, 4, 7	$[PtH_2O]^{2+}$, Met1, 4, 7	$[PtOH]^+$, Met1, 4, 7	trans-$[Pt(NH_3)_2]^{2+}$, Met1, 4	trans $[Pt(NH_3)_2]^{2+}$, Met1, 7	trans $[Pt(NH_3)_2]^{2+}$, Met1, 7
No QM atoms	36	38	40	39	33	33	33
No capping Hydrogens	3	3	3	3	2	2	2
Box sizes (Å) (x,y,z)	(42.1, 42.1, 44.7)	(43.2, 48.2, 46.4)	(43.4, 43.4, 47.1)	(42.6, 42.6, 46.3)	(47.6, 48.7, 46.5)	(46.3, 48.5, 44.9)	(49.1, 48.2, 45.9)

Table 4.1 Selected features of the simulated models

4.2.2 QM/MM simulations

The six representative structures of C1–C3 and T1–T3 underwent hybrid Car-Parrinello QM/MM simulations [167, 168, 169, 170, 171]. We used the CPMD program [188] combined with the classical MD GROMOS96 code [189], through the interface developed by Rothlisberger *et al.* [168, 169]. The MM regions comprised the peptide frame, the solvent, and the counterions (Fig. 4.1). These were described with the AMBER parm99SB force field [173] with the "Stony Brook" modification for the backbone torsions [190], the TIP3P model for the solvent [192] and the Smith&Dang force field for the counterions [193]. The QM part concerns the platinum coordination polyhedron (Fig. 4.1) and was treated by density functional theory (DFT), with the BLYP recipe for the exchange-correlation functional [195, 196]. The wavefunction was expanded in a plane-wave basis set up to an energy cutoff of 70 Ry. Only the valence electrons were treated explicitly (in the case of Pt, electrons in the n=5 shell are also included in the valence). The core electrons were described using norm-conserving pseudopotentials of the Martins-Troullier type [197]. Isolated system conditions were applied [198]. The β carbon (CB) atoms linked to the MM part by covalent bonds were saturated by "capping" hydrogen atoms[170]. The electrostatic interactions between QM and MM atoms were calculated using a fully Hamiltonian hierarchical scheme [169]. In particular, the short range electrostatic interactions between QM and MM part were taken explicitly into account within a radius of 5.3 Å around every QM atom using an appropriate modified Coulomb potential to prevent electron spill-out [169]. Beyond this first shell and within 10.6 Å, the electrostatic interactions are calculated using the D-RESP charge scheme [168]. In the outermost region, a multipole expansion scheme is used [168]. A time step of 0.097 fs was used. A fictitious electronic mass of 400 a.u. for the Car-Parrinello dynamics was employed. Constant temperature conditions were achieved using the Nose-Hoover chain thermostat[199, 200, 201]. Nonbonded interactions were evaluated as in the previous MD simulations. 2000 steps of simulated annealing were performed to relax the representative snapshots of the previous MD simulations. Then, the systems were slowly heated up to 300 K by

gradually increasing the target temperature (30 K after every 1000 steps). Finally, 25 ps-long QM/MM simulations were carried out.

4.2.3 Force matching procedure

AMBER-type force field parameters of the QM regions were obtained by applying the force matching procedure of Refs. [174, 175] and described in section 3.1.6 to 300 QM/MM conformations, selected at every 0.08 ps. The atomic partial charges (Tables. A.2–A.3 in Appendix) were obtained through a fit to the electrostatic potential and the electric field. These, along with van der Waals parameters taken from AMBER parm99SB force field [173, 190], were subtracted from the total forces acting on the QM atoms to yield the bonded forces. A least-squares fit of the latter allowed to obtain the AMBER-type bond, angle, and dihedral angle force field parameters (Tables. A.4–A.6 in Appendix). Equilibrium bond lengths and bond angles were taken from the QM/MM averaged values. The parameters for Met, water, and OH turned out to differ only by 30 to 40% relative to the original AMBER parameters.

4.2.4 Replica exchange with solute tempering (REST) simulations

Apo, C1–C3, and T1–T3 underwent, respectively, 0.5, 0.2, and 0.2 μs replica exchange with solute tempering (REST) simulations [176] in the NPT ensemble (see Sec. 3.1.9, for more details). The initial structures of the platinated peptides were the last QM/MM snapshots, whilst for Apo, the initial linear structure inserted in water was used. Eight replicas covered a temperature range from 300 K to 600 K for each system (Fig. 4.2). The force field was the same as that of the previous MD simulations, except that the platinated moieties were described by the parameters obtained via the force matching procedure.

First, each replica was prepared independently by 5-ns-long MD simulations, then the REST simulations were performed with the replica exchange being attempted after every 5 ps. The acceptance ratios were between 15 and 25%. The integration step of 2 fs was used. In each replica, the temperature was controlled by Nose-Hoover chain thermostat [199, 200, 201], whereas the pressure was controlled by Parrinello-Rahman barostat [202]. The REST simulations were carried out with the Gromacs 4.5.3 program [203]. The DSSP approach [204, 205] was used to assign the secondary structures to the conformations along each of the REST trajectories. The percentage of each secondary structure type was then estimated from the frequency of its occurrence during the REST simulations.

Next, 25 ps-long QM/MM simulations were carried out for Apo starting from the representative structures obtained from REST simulation. The QM part of the Apo model includes 3 Met residues (Fig. 4.1). The MM part consists of the peptide frame, the solvent, and the counterions. Exactly the same procedure as the one carried out for the platinated peptides (see Sec. 4.2.2) was employed. Finally, the

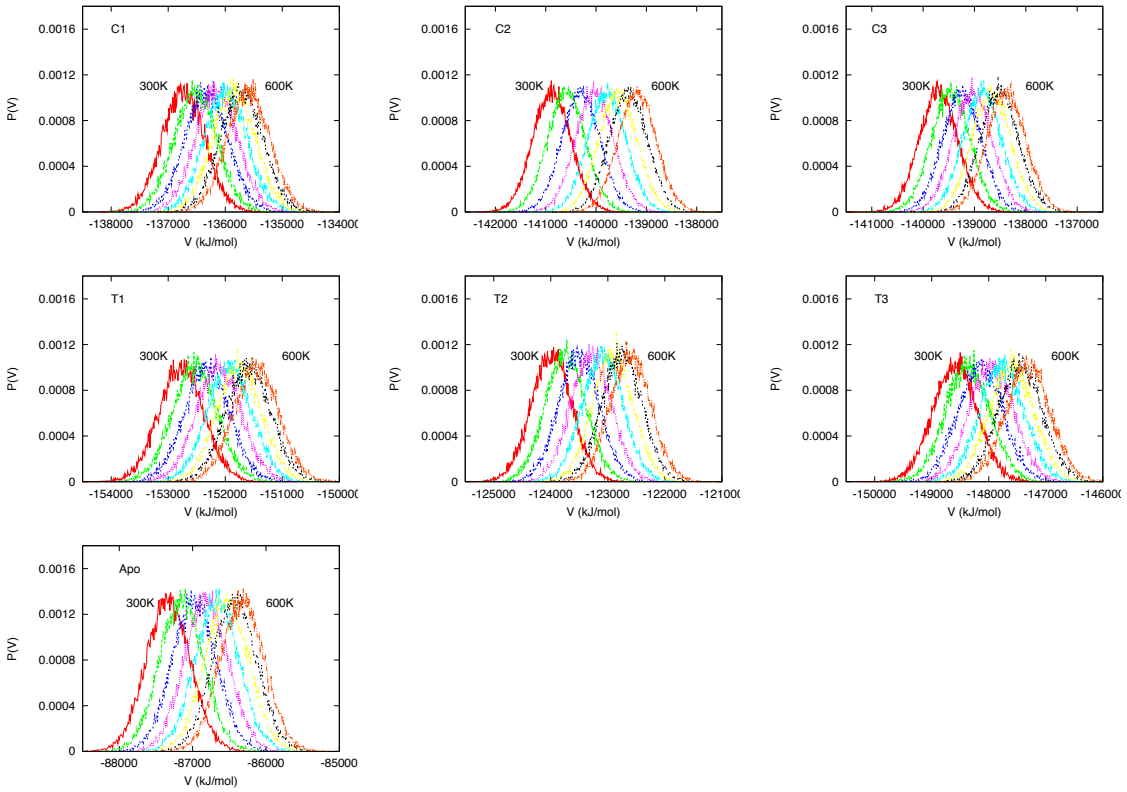


Figure 4.2 Potential energy distribution of 8 replicas (in different colors) simulated with REST method for the platinated peptides C1-C3 and T1-T3 and the apopeptide.

last snapshots of the REST trajectories of Apo, C1–C3 and T1–T3 underwent 25-ps-long QM/MM simulations with the same QM/MM procedure described in Sec. 4.2.2.

4.2.5 Computational spectroscopy

NMR chemical shifts and EXAFS spectra were evaluated based on the QM/MM trajectories of Apo and the platinated peptides before and after the REST simulation. The CD spectra were calculated from the REST trajectories.

4.2.5.1 CD

The CD spectra were averaged values over 500 equidistant snapshots along the 200-ns-long REST trajectories. The web interface DichroCalc [206, 207], which implements the theoretical approach described in Sec. 3.2.3, was used. The CD spectra turned out not to change with increasing REST simulation time (Fig. 4.3), suggesting that these calculated properties are well converged during the simulations.

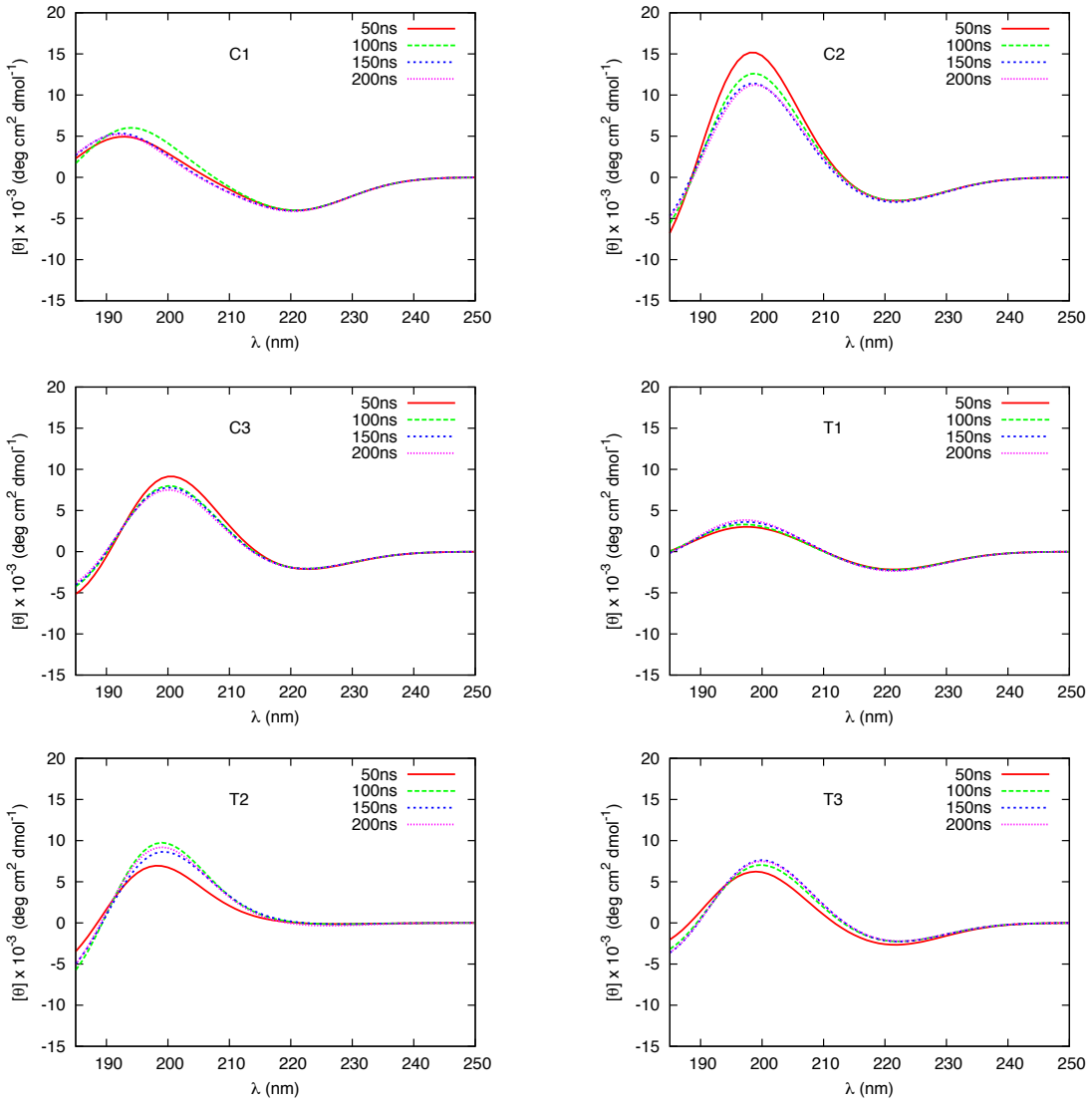


Figure 4.3 Running averages of CD spectra calculated from the REST trajectories for cisplatn-*Mets7* (C1-C3) and transplatin-*Mets7* (T1-T3).

4.2.5.2 NMR

^{1}H, ^{13}C, and ^{195}Pt NMR shielding constants were calculated as averages over the 50 equidistant QM/MM snapshots. For the first two nuclei (^{1}H and ^{13}C), we used the approach in Refs. [208] and [209] for the QM parts of C1–C3, T1–T3, and Apo. Because the QM/MM boundary cuts through the CA-CB bonds, the CB and HB atoms are very close to the MM part. They are expected to be overpolarized by the nearby MM charges. Hence, the calculation of shielding constants was not carried out for them. The shielding constants were converted to the NMR chemical shifts using the shielding constant of tetramethylsilane (TMS). See Appendix A.2 for details on conversion of the shielding constant to the chemical shifts.

For ^{195}Pt NMR shielding constants we used the ADF code [211, 212, 213] on the QM parts of C1–C3 and of T1–T3. The quadruple-quadruply polarized basis set (QZ4P) for the Pt and the triple-singly polarized basis set (TZP) for the other atoms were

used. The calculations were carried out at the DFT-PBE level [214]. This choice of basis sets and functional was shown to give reliable results for the calculation of ^{195}Pt chemical shifts [215]. Scalar relativistic effects were taken into account with the zero order regular approximation (ZORA) approach [216, 217]. The MM atoms were included as point charges. The chemical shift was referenced to that of $[PtCl_6]^{2-}$ in water at room temperature, as obtained by an additional QM/MM simulation (see Appendix A.3).

4.2.5.3 EXAFS

The EXAFS spectra were calculated at the Pt edge. The FEFF6L code [218, 219, 220] was applied over 500 equidistant QM/MM snapshots (See Sec. 3.2.2 for the theoretical approach). The FEFF calculation of EXAFS spectra are cheaper than that of NMR chemical shifts. Therefore, a larger number of QM/MM snapshots were used for EXAFS calculations. For each snapshot, a cluster of atoms within 5 Å from the platinum atom was extracted, taking into account the periodic boundary conditions, and input into the FEFF6L program to generate the EXAFS spectrum. The k^3-weighted EXAFS spectra were Fourier-transformed using the IFEFFIT program [221] to obtain the real-space representation.

4.3 EXAFS measurements

EXAFS measurements were done by Arnesano, F.; Scintilla, S. and Natile, G. [63].

The samples were prepared as follows. 3.0 mM aqueous solutions of Apo containing 100 mM NaCl, were treated with one mol equivalent of either cisplatin or transplatin and the mixtures were kept for 24 h at 310 K before recording the XAS spectra. The adduct formation between Mets7 and the selected platinum complex was checked by ESI-MS as previously described [142]. The samples were loaded in Teflon cells with Kapton windows, mounted in a two-stage Displex cryostat, and kept at 20 K during the data collection. XAS spectra were recorded at beamline D2 of the European Molecular Biology Laboratory (EMBL) outstation at Deutsches Elektronen Synchrotron (DESY, Hamburg, Germany). The DORIS storage ring operated at 4.5 GeV with the positron beam current ranging from 150 to 95 mA. Ionization chambers in front and behind the sample were used to monitor the beam intensity. A Si(111) double-crystal monochromator scanned X-ray energies around the platinum L3-edge (11564 eV). The X-ray absorption spectra were recorded as platinum L3 fluorescence spectra with a Canberra 13-element germanium solid-state detector. For both samples, 10 scans were collected and averaged to ensure comparable statistics. Data reduction, such as background removal, normalization, and extraction of the fine structure, was performed with KEMP [222]. The extracted platinum L3-edge extended X-ray absorption fine structure (EXAFS) data were converted to photoelectron wave vector k-space and weighted by k^3. The full, k^3 weighted, EXAFS spectra (30-1000 eV above E_0) and their Fourier transforms (FT) calculated over the range 2.0-16.0 Å^{-1} were compared with theoretical simulations.

	C1	C2	C3	T1	T2	T3
Averaged RMSD (Å)	1.4	1.1	1.6	1.8	2.2	2.0

Table 4.2 The averaged RMSD of the REST structures with respect to the corresponding last QM/MM structures.

4.4 Results and discussion

In the cisplatin-*Mets7* adducts (C1–C3, Fig. 4.1), the three methionine sulfurs act as donor atom to the Pt(II) ion, while the remaining fourth position may be occupied by chloride (C1), or water molecule (C2), or hydroxyl group (C3). In the transplatin-*Mets7* adducts (T1–T3), the two ammine ligands remain bound to Pt in *trans* arrangement. The other two ligands come from the two methionine sulfurs of either Met1,4 (T1), Met1,7 (T2) or Met4,7 (T3). We used QM/MM simulations and replica exchange with solute tempering (REST) method to investigate the Pt(II) ion coordination and the conformational properties of these platinated peptides, along with the conformational feature of the unbound peptide (Apo, Fig. 4.1). The direct comparison between the calculated and experimental spectroscopic data (EXAFS from this work, along with NMR and CD from Ref. [142]) established the accuracy of our structural predictions.

4.4.1 Peptide conformation

During the QM/MM simulations, the peptides' conformations do not change significantly (RMSD between the first and the last snapshot < 0.6 Å). In addition, the structures obtained from REST simulations are not too dissimilar to those from QM/MM simulations (Fig. 4.1): The averaged RMSD of REST structures with respect to the last QM/MM snapshots ranges from 1.1 to 2.2 Å (Tab. 4.2). The C1–C3 models are less flexible than the T1–T3 ones, as indicated by an analysis of the RMSF values. These are less than 1 Å for the first and between 1.1 and 1.5 Å for the latter (Tab. 4.3).

The content of the secondary structure (ss) during the REST simulations is given in Tab. 4.4. In the case of cisplatin-*Mets7* adducts, C2 and C3 contain more β-turn (>47%) than coil (<38%). In contrast, C1 contains more coil (50%) than β-turn (30%), and a small content of α-helix (8%). In the case of transplatin-*Mets7*, T1 contains mostly coil (59%) and only 26% of β-turn. Instead, both T2 and T3 have more β-turn than coil. Bends are found in all cases with rather small content ($<$ 16%). In the models with large β-turn contents, the latter is more likely to be found between the platinated methionines. The large content of turn appears to correlate with the emerging β-turn-type band in the calculated CD spectrum (Fig. 4.4).

The calculated CD spectrum of C3 is in good agreement with the experimental one (Fig. 4.4). It has a positive peak at 200 nm, which features the β-turn conformation [223]. The CD spectrum of C2 model has a similar shape but its intensity is slightly larger than that of the experimental spectrum. In contrast, in the spectrum of C1, the positive peak is shifted towards a lower wavelength (190 nm) compared to the

	C1	C2	C3	T1	T2	T3
RMSF (Å)	0.64	0.45	0.77	1.30	1.13	1.44

Table 4.3 The root mean square fluctuations (RMSF), which is defined as the standard deviation of the backbone RMSD. The smaller the value of RMSF, the more rigid the conformation, and vice versa.

	Coil	Bend	Turn	Helix
C1	50%	12%	30%	8%
C2	38%	15%	47%	0%
C3	37%	14%	49%	0%
T1	59%	15%	26%	0%
T2	39%	13%	48%	0%
T3	42%	11%	47%	0%

Table 4.4 DSSP secondary structure assignment for the ensembles of conformations generated by REST simulations.

experimental peak. In addition, it features a weak negative band at 210 nm. Hence, C2 and C3 reproduce better the β-turn feature of the experimental CD spectrum than C1. Since the only difference between C2 and C3 is the state of deprotonation of the coordinated solvent molecule (Fig. 4.1), an equilibrium between the two species may be present. The calculated CD spectrum of T1 reproduces fairly well the experimental one (Fig. 4.4). Those of T2 and T3 have a very large positive band at about 200 nm, which features their β-turn conformation. This contrasts with the experimental observation that the binding of transplatin to *Mets7* only slightly modifies the random coil structure of the free peptide [142]. This might imply that only T1 gives an important contribution to the total CD spectrum of the transplatin-*Mets7* adduct, while T2 and T3 are almost irrelevant.

The calculated CD spectrum of Apo has a positive band at 195 nm and a negative band at ∼210nm, which suggest some content of beta sheet structure. However, the experimental spectrum is characteristic of a random coil conformation [142]. The reason behind the discrepancy between experiment and theory might be due to sampling issues and to limitation of the force field. In particular, simulations based on the AMBER force field[173, 190], which has been used here, might bias random coiled structures towards the formation of secondary structure elements [226, 227, 228, 229]. In this respect, we stress here that the main focus of the paper is to reproduce the structural determinants of the platinated peptides, for which our computational procedure appears to be relatively accurate.

4.4.2 Platinum coordination

In the QM/MM simulations, the Pt geometry deviates only slightly from the ideal square-planar geometry, e.g. the difference from a square angle is less than 10% (Tab. 4.5). The Pt-S bond lengths range from 2.3 to 2.4 Å (Tab. 4.5). These

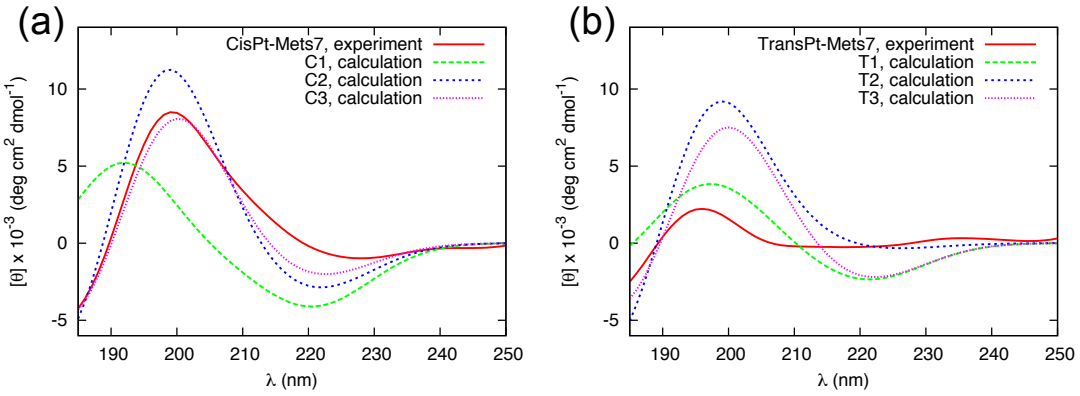


Figure 4.4 Calculated and experimentally measured [142] CD spectra of cisplatin-*Mets7* (a) and transplatin-*Mets7* (b) adducts.

values are very similar to those of corresponding model compounds ($[Pt(SH_2)_3Cl]^+$, $[Pt(SH_2)_3H_2O]^{2+}$, $[Pt(SH_2)_3OH]^+$ and *trans*-$[Pt(SH_2)_2(NH_3)_2]^{2+}$) optimized in the gas phase (Tab. A.1 in Appendix). Oxygen atoms from water molecules and/or the peptide terminus occupy the axial positions of the platinum complex, at a distance from the metal ranging from 2.6 to 2.8 Å (Fig. 4.5). In the REST simulations, the bond lengths and angles are very similar to those of QM/MM simulations (Tab. 4.5).

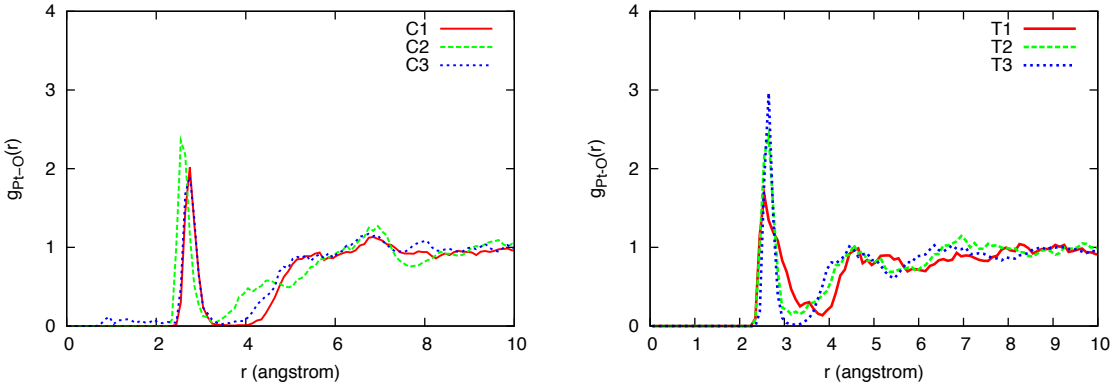


Figure 4.5 Pt-O radial pair distribution functions calculated from the QM/MM trajectories. Left: cisplatin-*Mets7*, right: transplatin-*Mets7*.

From the QM/MM calculations, local spectral properties can also be obtained. Despite the short time scale, the QM/MM calculations can provide rather accurate spectroscopic properties of the Pt site. This is because the overall conformations of the peptides, and in particular of the Pt site, in the QM/MM simulations are not too dissimilar to those obtained by the subsequent sampling-enhanced REST simulations. Furthermore, the QM region is relatively well sampled within the QM/MM time scale (see Sec. Appendix A.4). Here, we present EXAFS and NMR results from QM/MM simulations. The results of NMR obtained from QM/MM simulations after the REST simulations are reported here, whereas these berfore the REST simulations are reported in the Appendix.

Cisplatin/Mets7						
	C1		C2		C3	
	QM/MM	REST	QM/MM	REST	QM/MM	REST
Pt-S(Met1)	2.41 (0.07)	2.43 (0.07)	2.41 (0.06)	2.40 (0.06)	2.38 (0.06)	2.40 (0.07)
Pt-S(Met4)	2.39 (0.06)	2.40 (0.06)	2.34 (0.05)	2.35 (0.06)	2.40 (0.06)	2.39 (0.06)
Pt-S(Met7)	2.38 (0.05)	2.44 (0.07)	2.39 (0.05)	2.40 (0.06)	2.40 (0.07)	2.38 (0.07)
Pt-Cl	2.35 (0.05)	2.34 (0.05)				
Pt-O			2.12 (0.05)	2.14 (0.06)	2.04 (0.05)	2.06 (0.05)
O-H			1.00 (0.01)		0.99 (0.01)	
S(Met1)-Pt-S(Met4)	91.8 (3.2)	92.2 (4.4)	82.6 (3.1)	90.7 (3.5)	94.5 (3.1)	91.6 (4.0)
S(Met4)-Pt-S(Met7)	93.0 (3.1)	91.7 (3.4)	95.6 (3.2)	90.0 (3.3)	91.3 (3.3)	92.7 (3.7)
S(Met7)-Pt-Cl	90.0 (2.9)	88.5 (4.8)				
S(Met7)-Pt-O			92.8 (4.6)	90.1 (4.0)	86.4 (3.6)	87.6 (4.1)
Cl-Pt-S(Met1)	85.6 (3.4)	87.7 (5.2)				
O-Pt-S(Met1)			89.3 (4.3)	89.2 (4.1)	87.9 (3.7)	88.1 (4.4)
Transplatin/Mets7						
	T1		T2		T3	
	QM/MM	REST	QM/MM	REST	QM/MM	REST
Pt-S(Met1)	2.38 (0.06)	2.39 (0.07)	2.38 (0.06)	2.35 (0.06)		
Pt-S(Met4)	2.39 (0.06)	2.38 (0.07)			2.39 (0.06)	2.39 (0.06)
Pt-S(Met7)			2.39 (0.06)	2.35 (0.07)	2.39 (0.06)	2.39 (0.07)
Pt-N(Am1)	2.09 (0.05)	2.08 (0.05)	2.09 (0.04)	2.07 (0.05)	2.10 (0.05)	2.09 (0.05)
Pt-N(Am2)	2.11 (0.05)	2.08 (0.05)	2.10 (0.04)	2.07 (0.05)	2.09 (0.04)	2.09 (0.05)
S(Met1)-Pt-N(Am1)	94.5 (3.2)	89.6 (3.5)	95.5 (3.5)	89.7 (3.6)		
S(Met1)-Pt-N(Am2)	84.4 (3.1)	90.2 (3.3)	85.1 (3.2)	89.7 (3.6)		
S(Met4)-Pt-N(Am1)	84.7 (3.1)	90.2 (3.4)			95.2 (3.5)	90.1 (3.5)
S(Met4)-Pt-N(Am2)	97.3 (3.4)	90.0 (3.5)			85.1 (3.3)	90.0 (3.4)
S(Met7)-Pt-N(Am1)			83.3 (2.9)	90.3 (3.8)	93.6 (3.1)	90.0 (3.5)
S(Met7)-Pt-N(Am2)			95.9 (3.1)	90.2 (3.9)	86.2 (3.0)	90.0 (3.3)

Table 4.5 Structural parameters calculated from QM/MM and REST simulations for cisplatin-*Mets7* (C1-C3) and transplatin-*Mets7* (T1-T3) adducts. Standard deviations are shown in parentheses. Bond lengths are in angstrom and angles in degree.

4.4.2.1 EXAFS spectra

The calculated k^3-weighted EXAFS spectra at the Pt edge of C1–C3 and T1–T3 reproduce well the experimental data (Fig 4.6 a-b). Also the corresponding Fourier transforms (FT's) of the calculated EXAFS spectra are in fair agreement with those of the experiments (Fig. 4.6 c-d).

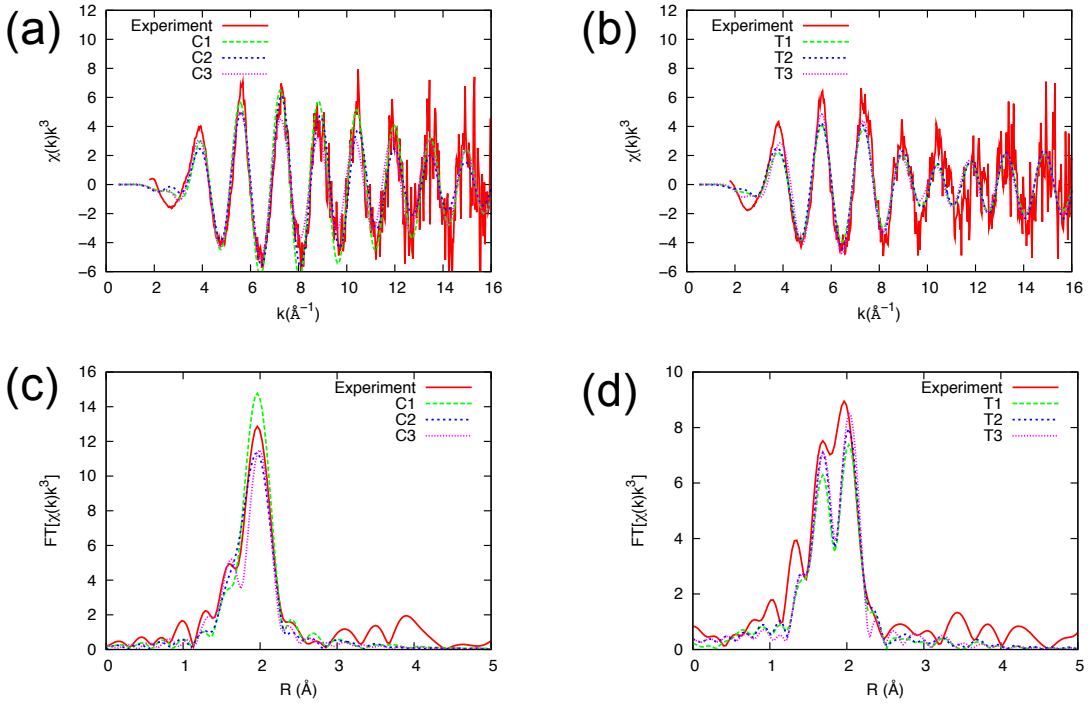


Figure 4.6 Calculated and experimental EXAFS spectra at the Pt edge of cisplatin-*Mets7* (a) and transplatin-*Mets7* (b) adducts. The experimental spectrum has been measured in this work. Fourier Transforms of the spectra in (a) and (b) are given in (c) and (d), respectively. They are plotted as a function of the scattering path length of the photoelectron R. Note that R does not correspond to a real interatomic distance since the peak positions are shifted inward, with respect to the bond distances; due to the phase shifts of the photoelectron (i.e. increase in its kinetic energy caused by the electronic potentials of both the scattering and the absorbing atoms) [225]. The shift, therefore, depends on both the absorbing and scattering atom types, and in our cases is approximately 0.4 Å (Tab. 4.5.

The FEFF approach of EXAFS calculation is based on the multiple-scattering theory, in which the total EXAFS are the sum over all possible scattering paths of the photoelectron [220]. It is, therefore, possible to separate contributions due to each type of ligand's atoms and to assign peaks of EXFAS to the corresponding atoms. In Fig. 4.7, EXAFS is calculated separately for each type of donor atoms in order to distinguish between peaks due to different atoms.

For the cisplatin-*Mets7* adducts (Fig. 4.6 c) the three FT's feature a large peak, mainly caused by the back-scattering from the three Pt-bound sulfur atoms and a chloride (C1) or a water oxygen (C2); a small side peak on the left which appears

clearly only in C3 is due to the hydroxyl oxygen (Fig. 4.6 c). Oxygen atoms at the axial position are found to give small contribution to the EXAFS signals of all models. The main contributions from each type of ligand's atoms to the total EXAFS spectra are shown in (Fig. 4.7). The peak heights of C2 and C3 are lower than the experimental one. On the contrary, the peak height of C1 is greater. We note that the EXAFS spectra were taken at high chloride concentration (100 mM, see Methods), while the CD spectra were taken at low chloride concentration [142]. Therefore model C1, which did not appear to give significant contribution to the CD spectrum, could give a significant contribution to the EXAFS spectrum.

For the transplatin-*Mets7* adducts (Fig. 4.6 d), the FT's feature two large distinct peaks in the first coordination shell. The first peak is due to the ammine nitrogen atoms and the second peak is due to the Met sulfur atoms (Fig. 4.7). The large difference in length between Pt-S and Pt-N bonds (Tab. 4.5) causes such a splitting. Again, oxygen atoms at the axial position are found to give small contribution to the EXAFS signals of all models The magnitude of FT's of T1–T3 is slightly smaller than those obtained by experiment.

4.4.2.2 NMR chemical shifts

The calculated ^{13}C and ^{1}H NMR chemical shifts of the platinated and Apo peptides reproduce the experimental values within the indicated standard deviations (Tab. 4.6). In particular, as observed for the experimental data, the values of ^{1}H and ^{13}C chemical shifts at the γ-CH_2 and ϵ-CH_3 residues of the platinated peptides (C1–C3 and T1–T3) are larger than those of the Apo (Tab. 4.6). This is consistent with the experimental observation that the binding of cisplatin and transplatin to the *Mets7* peptide causes a large downfield shift in the NMR signals for the γ-CH_2 and ϵCH_3 atoms of all three methionines [142]. The downfield shift of γ-CH_2 and ϵ-CH_3 atoms of all three methionines, also in the case of the transplatin-*Mets7* adducts where only two methionines are coordinated to platinum, indicates that the three models are in fast exchange even if they can contribute differently to the overall spectrum.

The calculated ^{195}Pt chemical shifts of C1 and C3 reproduce quite well the experimental data (Tab. 4.7). That of C2 deviates significantly from the experimental value. This might imply that the C2 model is less relevant than C1 and C3. As already pointed out in the discussion of the EXAFS spectra, also the NMR spectra were taken at a rather high chloride concentration (3.0 mM). Under these conditions model C1 could gain relevance and contribute to the overall NMR spectra. The ^{195}Pt NMR chemical shifts of T1–T3 do not differ very much. The difference between cisplatin-*Mets7* and transplatin-*Mets7* ($\sim$1000 ppm) is fairly well reproduced by our calculations.

The use of the REST trajectories for calculating the CD spectra, whereas the QM/MM trajectories have been used to calculate the NMR and EXAFS spectra, deserve an additional comment. The CD spectrum is obviously sensitive to the global conformations of the peptides' backbone. Therefore, we used for its calculation a sampling-enhanced technique, such as REST [176], that samples extensively the conformational space. On the other hand, the NMR chemical shifts of the Pt ligands and the EXAFS spectra at the Pt edge are expected to be sensitive to the local geometry (bond lengths and bond angles) of the plati-num coordination shell.

^{13}C NMR											
		Calculated values							Experimental values		
Met#	Atom name	Apo	C1	C2	C3	T1	T2	T3	Apo	cisplatin /Mets7	transplatin /Mets7
Met1	C_γ	31.1 (3.5)	35.7 (5.8)	37.2 (5.2)	39.3 (4.9)	33.9 (4.1)	39.5 (4.9)		29.3	34.1	34.3
	C_ε	17.3 (2.3)	21.3 (4.9)	21.6 (3.0)	23.3 (2.8)	22.2 (2.7)	20.5 (5.1)		14.2	19.9	18.5
Met4	C_γ	30.4 (2.4)	40.4 (5.2)	40.3 (5.6)	39.2 (5.2)	37.9 (4.3)		38.8 (5.6)	29.3	34.1	34.3
	C_ε	17.2 (2.5)	18.5 (4.0)	21.5 (3.2)	21.6 (2.1)	15.9 (3.1)		21.2 (3.8)	14.2	19.9	18.5
Met7	C_γ	32.3 (2.6)	39.5 (5.9)	39.6 (5.2)	37.2 (5.4)		40.2 (5.3)	39.1 (5.8)	29.3	34.1	34.3
	C_ε	16.9 (2.7)	17.5 (4.1)	18.0 (4.2)	15.8 (3.7)		21.1 (4.5)	21.0 (3.4)	14.2	19.9	18.5
^{1}H NMR											
		Calculated values							Experimental values		
Met1	H_γ	2.65 (0.45)	2.78 (0.36)	2.84 (0.35)	3.18 (0.44)	3.52 (0.34)	3.12 (0.36)		2.51	2.89	3.03
	H_ε	2.41 (0.26)	2.60 (0.43)	2.62 (0.42)	2.71 (0.41)	2.76 (0.27)	2.79 (0.42)		2.03	2.35	2.51
Met4	H_γ	2.70 (0.40)	2.90 (0.51)	3.05 (0.53)	3.02 (0.39)	3.47 (0.27)		3.61 (0.45)	2.51	2.89	3.03
	H_ε	2.24 (0.37)	2.61 (0.50)	2.57 (0.52)	2.44 (0.49)	2.54 (0.29)		2.62 (0.30)	2.03	2.35	2.51
Met7	H_γ	2.64 (0.34)	3.36 (0.59)	3.21 (0.53)	3.03 (0.52)		3.68 (0.45)	3.32 (0.40)	2.51	2.89	3.03
	H_ε	2.45 (0.38)	2.79 (0.59)	2.66 (0.37)	2.63 (0.61)		2.56 (0.33)	2.62 (0.38)	2.03	2.35	2.51

Table 4.6 ^{13}C (top) and ^{1}H (bottom) NMR chemical shifts (in ppm) of the Pt-coordinated Met residues of cisplatin-*Mets7* (C1–C3) and transplatin-*Mets7* (T1–T3) adducts after REST simulations. The corresponding values before REST are reported in Tables. A.8–A.10 in Appendix. Standard deviations are shown in parentheses. See Fig. 4.1 for atom name conventions.

Cisplatin-Mets7			
Experiment	-4100	Experiment	-2900
C1	- 3815 (321)	T1	- 2793 (353)
C2	- 2793 (361)	T2	- 3121 (347)
C3	- 3571 (382)	T3	- 2751 (286)

Table 4.7 ^{195}Pt NMR chemical shifts (in ppm) of the cisplatin-*Mets7* (C1–C3) and transplatin-*Mets7* (T1–T3) adducts after REST simulations. The corresponding values before REST are reported in Tab. A.11 in Appendix. Standard deviations are shown in parentheses.

QM/MM describes accurately the metal ion coordination polyhedra in metal-based biomolecules [170]. Therefore, we can expect accurate results based on the QM/MM trajectories. However, in this case sampling is obviously much more limited than in the case of the REST simulations. To address this issue, we compare the results from two sets of QM/MM calculations, before (Tables A.8–A.11, in Appendix) and after (Tables 4.6–4.7) the REST simulations. It turns out that the results are essentially the same, in spite of the fact that the conformations of the peptide frame at the end of the REST trajectories are rather different from those at the beginning. Based on these findings, we suggest that the EXAFS and NMR shifts do not depend significantly on the peptides' global conformations but rather they depend on the local environment. This feature has been found also in several other bioinorganic systems [230, 231, 232].

4.5 Conclusions

The cellular uptake of cisplatin is mediated by the high-affinity copper transporter *Ctr1* [30, 178]. The 8-residue long peptide (*Mets7*) is able to mimic the extracellular Met-rich N-terminal motifs of yeast *Ctr1* [29]. The reaction of cisplatin and transplatin with *Mets7* was characterized experimentally by Arnesano F. *et al.* [29]. Here, the structural and spectroscopic properties of cisplatin and transplatin adducts with *Mets7* (models C1–C3 and T1–T3, respectively, see Fig. 4.1) have been investigated, for the first time, by hybrid Car-Parrinello QM/MM simulations. Information about the overall conformation of the peptide could be deduced from CD spectra. To this end, 200 ns-long REST simulations were carried out for the adduct models. The simulations were based on the force field parameters obtained by force matching procedure [174] for the platinated moieties. The configurationally averaged CD spectra were calculated for each of the models.

It turns out that the CD spectra of different simulated species show rather different patterns. As for the CD spectrum of the cisplatin-*Met7* adduct, we notice that the band wavelengths of the spectra of C2 and C3 models, characteristic of a β-turn structure, is in a good accord with the experimental one. The spectrum of C3 matches the intensity of the experimental one better than that of C2. In contrast, the C1 model shows much worse correspondence, particularly in terms of band wavelengths. We conclude that (i) the spectrum of the C3 model, in which the ligands of the Pt(II) ion are three Met residues and a hydroxyl group, gives the most important contribution to the total CD. In addition (ii), in the experimental conditions of Ref. [29], the coordinated water molecule is deprotonated so as to reduce the net positive charge of the complex. In the case of the transplatin-*Mets7* CD spectrum, the band intensities and wavelengths of the T1 model show good agreement with the experimental ones. Those of the T2 and T3 models are instead significantly different from the experimental ones and show a strong signal characteristic of a β-turn structure. We conclude that the T1 model (coordination to platinum of Met1 and Met4, Fig. 4.1) is likely to give a much greater contribution to the total CD than those of the other two models.

Concerning the Pt coordination, since experimental data for the bond lengths and angles are not available, the consistency between the Pt-coordination geometry of

the simulated models and the real ones has been checked by comparing the calculated Pt EXAFS spectra of C1–C3 and T1–T3 with the experimental ones reported here for the first time. The calculated EXAFS spectra of the three models show good agreement with the experimental spectra. Differently from the CD spectra, the EXAFS spectra were taken at considerably higher chloride concentration. Hence, model C1 (chloro species), which did not contribute appreciably to the CD spectra, can instead give a significant contribution to the EXAFS spectra. Calculations show that, indeed, this is the case.

The calculated ^{1}H and ^{13}C NMR chemical shifts are in rather good agreement with the experimental chemical shifts. The differences in the calculated ^{1}H and ^{13}C NMR chemical shifts between Apo and platinated peptides confirm that platinum binding causes a downfield shift of the NMR signals belonging to the Met residues that are bound to platinum as observed experimentally. The calculations also confirm a contribution of models T2 and T3 to the overall spectra to account for the downfield shift observed for γ-CH_2 and ϵ-CH_3 of all methionines of *Mets7*. The calculated ^{195}Pt chemical shifts also show that model C1 can give a relevant contribution to the experimental spectrum taken at a rather high chloride concentration. In contrast, model C2, whose computed spectrum significantly deviates from the experimental one, contributes much less.

The calculated ^{1}H and ^{13}C NMR chemical shifts are in rather good agreement with the experimental chemical shifts. The differences in the calculated ^{1}H and ^{13}C NMR chemical shifts between Apo and platinated peptides confirm that platinum binding causes a downfield shift of the NMR signals belonging to the Met residues that are bound to platinum as observed experimentally. The calculations also confirm a contribution of models T2 and T3 to the overall spectra to account for the downfield shift observed for γ-CH_2 and ϵ-CH_3 of all methionines of *Mets7*. The calculated ^{195}Pt chemical shifts also show that model C1 can give a relevant contribution to the experimental spectrum taken at a rather high chloride concentration. In contrast, model C2, whose computed spectrum significantly deviates from the experimental one, contributes much less.

Different spectroscopies, combined with quantum chemical calculations, provide valuable complementary insights in bioinorganic systems [233]. Here, we have calculated a variety of spectroscopic data (CD, EXAFS, and NMR), which are readily compared with experiments. The exhaustive comparison of the calculated spectroscopic data with the experimental ones has been proved useful. It allows us to make structural prediction of platinated *Mets7* species in aqueous solution.

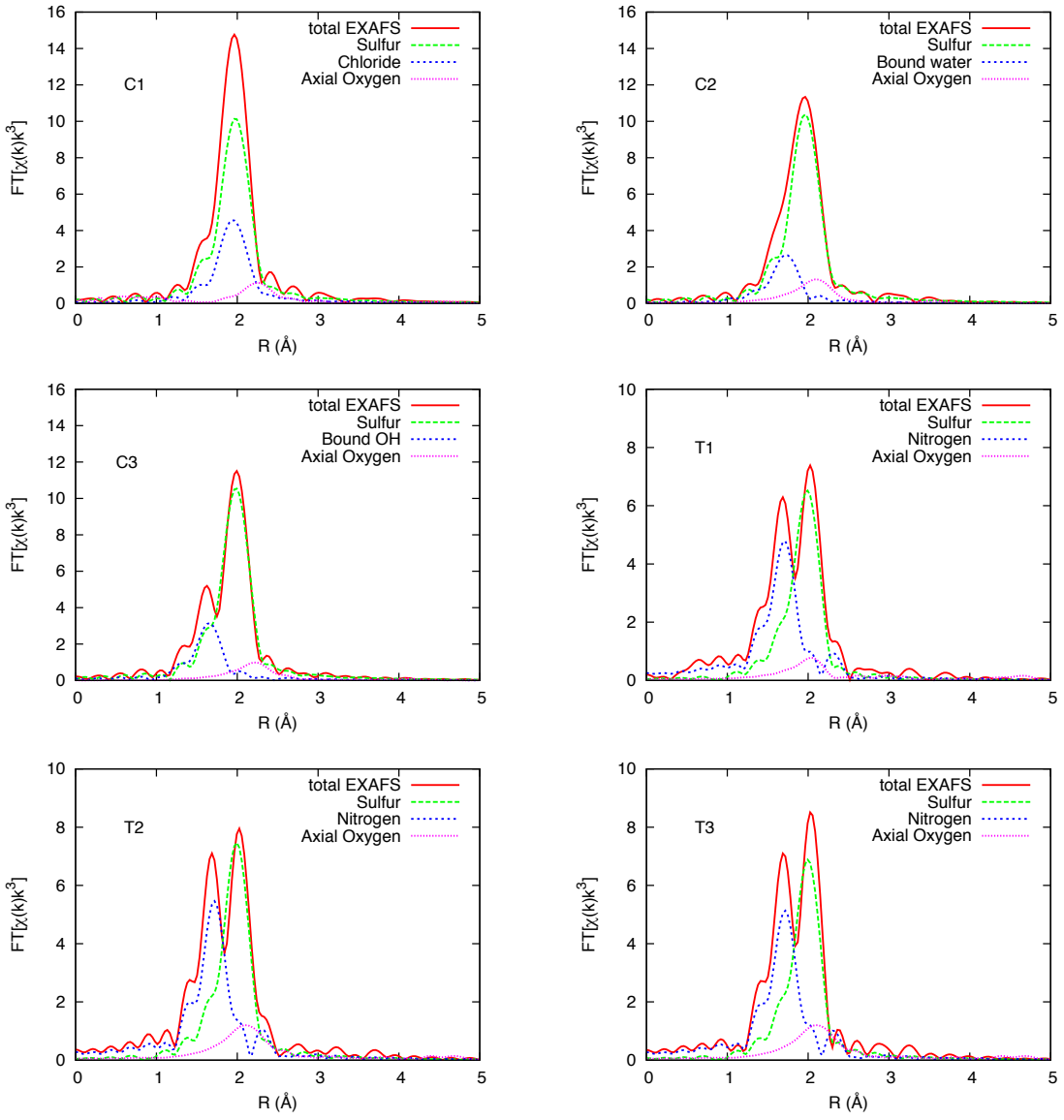


Figure 4.7 Contributions from different types of scattering atoms to the total EXAFS spectra, calculated from the QM/MM trajectories.

5

Binding of the platinated DNA to HMGB1 protein: preliminary results

5.1 Introduction

DNA is the primary target of cisplatin [1, 2, 4]. It binds preferentially to two adjacent guanine bases within a strand of DNA [1, 2, 4]. Cisplatin binding causes damages in DNA (cisplatin-DNA adducts) which, in turn, inhibit DNA replication [21, 22] and transcription [24, 25]. The inhibition of these vitally important cellular processes leads to apoptotic death of tumor cells.

Cisplatin-DNA adducts are recognized by many DNA-binding proteins [1, 3]. The most characterized among these proteins is the High Mobility Group Box 1 protein (*HMGB1*) [1, 24, 76, 78]. Interestingly, *HMGB1* has been found to inhibit Nucleotide Excision Repair (NER) of cisplatin-damaged DNA by binding to cisplatin-DNA adducts and shielding them from being repaired by NER proteins [72, 73, 74, 78] (see Sec. 2.2.2 for more details).

X-ray studies of cisplatin-d(CCUCTCTG*G*ACCTTCC)-d(GGAGAGACCTGGAAGG) in complex with the *HMGB1* domain A (G* is the platinated base) have shown the molecular 3-D structure of the cisplatin-DNA-*HMGB1* complex (Fig. 5.2) [76]. *HMGB1* consists of two DNA-binding domains (A and B), each containing three alpha helices forming an *L*-shape structure that facilitates DNA binding [77] (Fig. 5.1). The concave surfaces of *HMGB1* domain A form hydrogen bonds as well as hydrophobic and water-bridging contacts with the minor groove of cisplatin-DNA adduct. The binding affinity of *HMGB1* to cisplatin-DNA adduct at is high K_{d}[1] = 0.7 ± 0.1 nM [78].

[1] K_{d} is given by

$$K_{\mathrm{d}} = \frac{[\mathrm{P}][\mathrm{D}]}{[\mathrm{P} \cdot \mathrm{D}]}, \tag{5.1}$$

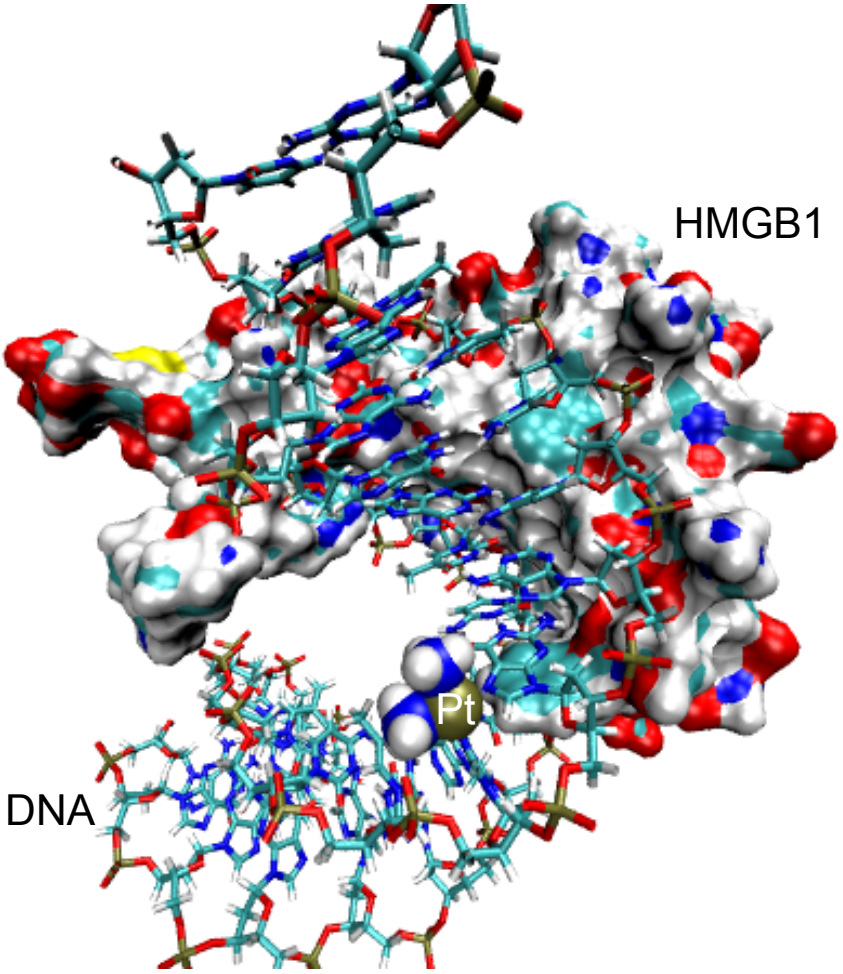


Figure 5.1 The X-ray structure of cisplatin-DNA-*HMGB1*. The DNA is shown with bond representation, HMGB1 with Connolly surface representation and cisplatin moiety in van der Waals balls. Hydrogen atoms missing in the X-ray structure were added using the Leap program [185].

Molecular recognition of cisplatin-DNA adducts by *HMGB1* protein is a key cellular event in inducing the antitumor effect of cisplatin [1, 3]. Therefore, determining the molecular mechanism of *HMGB1* binding to cisplatin-DNA adduct would be a next fundamental step towards better understanding of the molecular basis of action as well as resistance mechanism of cisplatin.

The free energy as a function of the collective variables (CVs) describing the association process of *HMGB1* and the platinated DNA is a key physical quantity of the binding process. Metadynamics method and its variants have emerged as a powerful technique to calculate such quantity [284, 285] (see Appendix A.5). It is based on a dimensional reduction and on a suitable history-dependent potential. First, it re-

where [P], [D] and [P·D] are concentrations of protein, DNA and protein-DNA complex, respectively. Since K_d is an equilibrium constant, it is related to the Gibbs free energy difference ΔG via

$$K_\mathrm{d} \propto \exp\left[-\frac{\Delta G}{k_B T}\right], \tag{5.2}$$

where k_B is the Boltzmann constant and T is the temperature.

quires identifying a set of CVs, which can properly describe the activated process of interest. The dynamics in the space of CVs is biased by a time-dependent potential, defined as a sum of Gaussians centered along the trajectory followed by the CVs [284]. This potential in time fills minima in the free energy surface. At the end, sum of the Gaussians provides the negative value of the free energy [284]. Metadynamics, as a variety of other free energy methods including thermodynamic integration and umbrella sampling, forces the biomolecules to explore their conformational space quite rapidly. This issue depends crucially on the choice of CVs, which must provide a satisfactory description of the reaction coordinates of the process we wish to investigate. If an important CV is not included, then the simulation will suffer from serious artifacts [284, 285]. This problem might be overcome by combining well-tempered metadynamics (a variant of metadynamics) with parallel tempering (PT) [286, 287, 288]. The strengths of both approaches are combined: Well-tempered metadynamics ensures the convergence of the free energy profile to a correct value, whereas PT exploits temperature to accelerate the sampling of all degrees of freedom including the CVs, which are not explicitly included in the metadynamics simulation. Therefore, large hysteresis, which is very commonly encountered in standard metadynamics when important CVs are neglected, can be avoided [287]. We plan to use this combined scheme to explore the free energy landscape of the cisplatin-DNA-*HMGB1* complex.

Here, we have performed a variety of calculations to derive force field parameters for the platinated moiety of DNA. We have considered the functional form of AMBER force field [173] in the parameterization procedure and we will use this for all the subsequent calculations. The AMBER force field is one of the most widely used force field for simulations of nucleic acids [289, 290]. Our calculations include a hybrid Car-Parrinello/molecular mechanics (QM/MM) simulation [167, 168, 169, 170, 171], force matching procedure [174, 175] and a 200-ns-long classical MD simulation.

In the next sections, we will present results from these calculations

5.2 Derivation of force field parameters for the platinated moiety

5.2.1 QM/MM simulation

The QM/MM simulation was based on the X-ray structure of cisplatin-d(CCUCTCT-G*G*ACCTTCC) -d(GGAGAGACCTGGAAGG) in complex wih *HMGB1* domain A (PDB code 1CKT) (Fig. 5.2) [76]. The complex were immersed in a box of water, Na^+ ions were added to neutralize the total charge (See Tab. 5.1 for details of the simulation setup). The whole system underwent 5ns-long preliminary classical MD simulation, in which the cisplatin moiety was kept fixed with SHAKE algorithm [235]. The Amber99SB-ILDN force field [173, 190, 191] was used for the protein and DNA. TIP3P model was used for waters [192] whereas Joung&Cheatham force field [194] was used for ions. The NAMD program [186] was used to carry out the MD simulation. The last snapshot was selected for the subsequent QM/MM simulation.

The CPMD program [188] combined with the classical MD Gromos96 code [189], through the interface developed by Rothlisberger *et al.* [168, 169] was used for the

No. of atoms	63937
No. of water molecules	20572
No. of Na^+ counterions	20
QM-Part	cis-[Pt-$(NH_3)_2$]$^{2+}$, G8, G9
No. of QM atoms	77
Box size (Å) (x, y, z)	(8.63, 8.63, 8.63)

Table 5.1 Selected features of the qmmm simulation box.

	QM/MM	Classical MD	X-ray
Pt coordination			
Pt-NE	2.078 (0.035)	2.079 (0.038)	1.98
Pt-NF	2.072 (0.040)	2.071 (0.039)	1.96
Pt-N7(G8)	2.073 (0.043)	2.074 (0.045)	1.89
Pt-N7(G9)	2.075 (0.052)	2.079 (0.049)	1.94
N7(G8)-Pt-N7(G9)	90.5 (2.6)	90.3 (2.9)	90.3
NE-Pt-N7(G8)	88.9 (3.3)	89.0 (3.1)	88.8
NE-Pt-N7(G9)	175.1 (2.5)	175.2 (2.6)	175.1
NF-Pt-N7(G8)	174.9 (2.7)	175.2 (2.6)	177.3
NF-Pt-N7(G9)	91.0 (3.0)	89.0 (3.2)	92.0
NF-Pt-NE	89.6 (3.0)	91.6 (3.3)	89.1
DNA parameters			
Total helical bend (°)	62.3 (5.6)	59.2 (10.6)	58.8
Rise G8-G9 (Å)	5.7 (0.3)	5.2 (0.3)	5.33
Tilt G8-G9 (°)	-27.4 (4.5)	22.0 (4.3)	-26.2
Roll G8-G9 (°)	69.3 (4.7)	62.0 (4.1)	69.1

Table 5.2 Comparison of the selected structural features of QM/MM, classical MD models with those of the X-ray structure. See Fig. 5.5 for the atom type convention.

QM/MM simulation. The QM part includes two guanine bases G8 and G9, the Pt^{2+} metal and two ammine (NH_3) group. The MM part consist of the rest of the simulation box (see Tab. 5.1 for more details). The QM/MM procedure is the same as that of the platinated *Mets7* in chapter 4.

Figure 5.3 shows the backbone RMSD of the QM/MM trajectory. There is no significant structural change during the 20-ps-long QM/MM simulation (RMSD between the first and the last snapshot $\sim$ 1.0 Å). The structural parameters of DNA such as helix bend, torsional angles were calculated by using the Web server Curves+ [292]. Table 5.2 shows comparison of the selected structural features of Pt-coordination and DNA between QM/MM and X-ray structures. The results obtained from QM/MM simulation do not differ significantly from the X-ray structure.

5.2.2 Force matching procedure

Next, the force matching technique of Refs. [174, 175] was applied to 200 conformations selected from the QM/MM trajectory to obtain the AMBER-type [173] force

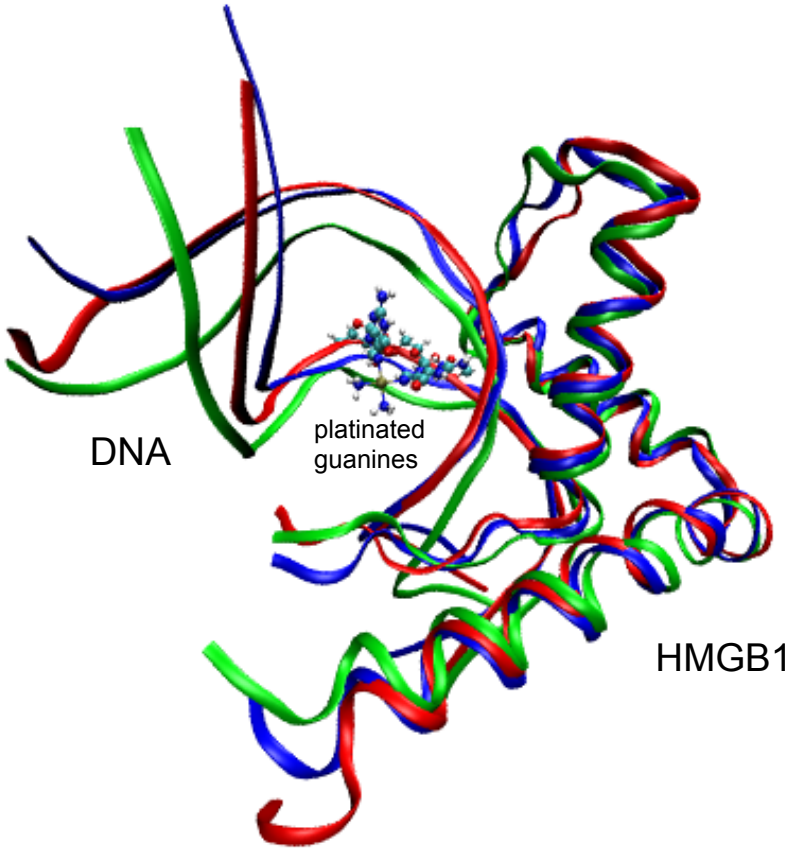


Figure 5.2 Complex of platinated DNA-*HMGB1*. The backbone conformations of X-ray structure (red), the QM/MM structure (blue) and the classical MD structure (green) are superimposed. Platinated guanines are shown in ball-and-stick.

field parameters for the plaitaned moiety. The procedure is essentially the same as that used for the platinated *Mets7* in chapter 4.

Partial charges are reported in Tab. 5.3. To minimize changes in original AMBER bonded parameters, only bonded parameters involved cisplatin moiety were parameterized. These parameters are reported in Tabs. 5.4–5.5.

5.3 Force field based MD simulation

The newly derived parameters together with the Amber99SB-ILDN force field [173, 190, 191] and the Barcelona modification [291] has be used to carry out a 200ns-long classical MD simulation for the cisplatin-DNA-*HMGB1* complex. The backbone RMSD is reported in Fig. 5.6 and some structural features in Tab. 5.2. The classical model reproduces well the Pt coordination geometry of the QM/MM simulation. Furthermore, during this 200ns-long classical MD simulation, the system is very stable (Fig. 5.6) and does not deviate significantly from the X-ray structure (the averaged RMSD with respect to the X-ray structure is $\sim$ 3Å). Most of the important contacts between the platinated DNA and *HMGB1* in the X-ray structure are preserved during the classical MD simulation (Fig. 5.7–5.8).

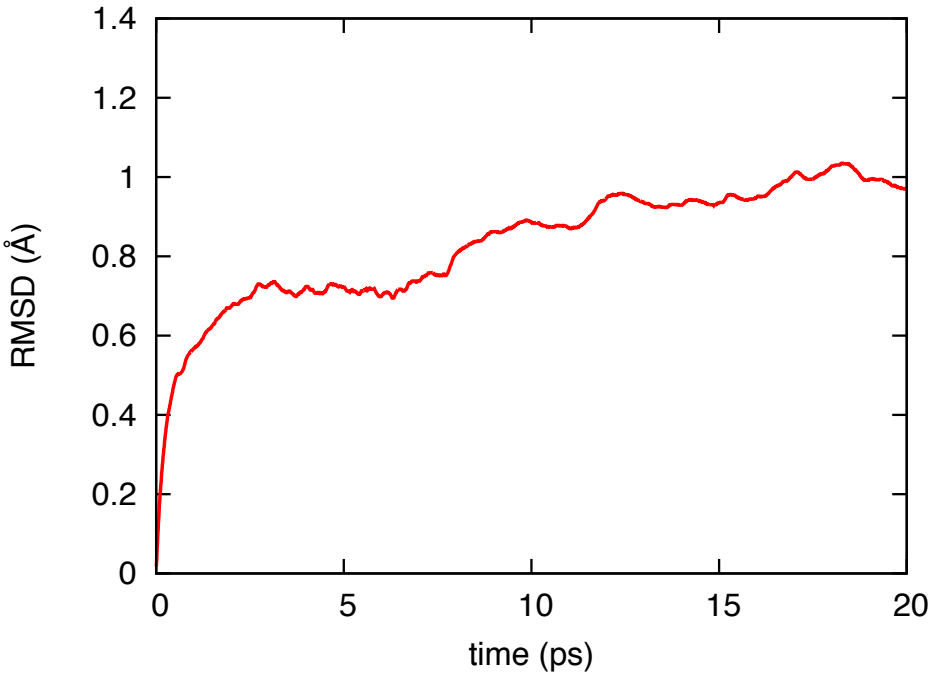


Figure 5.3 The backbone RMSD of QM/MM trajectory.

5.4 Summary

In summary, the AMBER-type force field parameters for the platinated site in the cisplatin-DNA-*HMGB1* complex has been derived using the force matching procedure [174, 175] for the QM/MM trajectory. Then the system has been equilibrated for 200ns with classical MD simulation based on the newly derived force field parameters for the platinated moiety and the Amber99SB-ILDN force field [173, 190, 191] with the Barcelona modification [291] for the rest of the complex. This equilibrated structure will be used for the subsequent metadynamics simulations.

Cisplatin moiety			
Atom names			
Pt	0.313426		
NE/NF	-0.636923		
HE/HF	0.389406		
Platinated Guanines			
	G8	G9	A10
P		1.075780	1.213865
OP		-0.833611	-0.839829
O5'		-0.418657	-0.300678
C5'		0.255192	-1.074532
H5'		0.039748	0.297395
C4'	-1.147331	0.169942	
H4'	0.273118	0.101920	
O4'	-0.583088	-0.692190	
C1'	1.727400	1.599090	
H1'	-0.293229	-0.215040	
N9	-0.569336	-0.340645	
C8	0.160867	-0.082580	
H8	0.329768	0.211014	
N7	-0.541038	-0.075359	
C5	0.435529	-0.061307	
C6	0.334949	0.742188	
O6	-0.545681	-0.658232	
N1	-0.561287	-0.709675	
H1	0.458696	0.478193	
C2	1.007432	1.138611	
N2	-0.919651	-1.214229	
H2	0.463393	0.542145	
N3	-0.844970	-0.632471	
C4	0.291287	0.282156	
C3'	0.939230	0.289710	
H3'	-0.188648	0.132338	
C2'	-0.127299	-0.991966	
H2'	-0.035253	0.270732	
O3'	-0.559364	-0.485245	

Table 5.3 Partial charges (in fraction of electronic charge) derived from the force matching procedure for the cisplatin-DNA-*HMGB1* complex. See Fig. 5.4 for the atom name convention.

Bonds		
	K_r	r_{eq}
Pt-N7	123	2.072
Pt-N6	116	2.075
Pt-N5	98	2.075
Pt-N4	102	2.073
N7-H7	364	1.04
N6-H6	382	1.04
Angles		
	K_θ	θ_{eq}
H7-N7-H7	31	109.5
H7-N7-Pt	26	109.5
H6-N6-H6	33	109.5
H6-N6-Pt	29	109.5
N7-Pt-N6	113	89.6
N7-Pt-N5	68	91.0
N7-Pt-N4	31	175.0
N6-Pt-N4	61	89.0
N6-Pt-N5	25	175.0
N4-Pt-N5	35	90.5
Pt-N5-CK	52	123.0
Pt-N5-CB	47	129.5
Pt-N4-CK	34	120.5
Pt-N4-CB	41	132.5

Table 5.4 The harmonic force constant K_r (kcal/mol $\mathring{A}^{-2}$), K_θ (kcal/mol rad^{-2}) of bond stretching and angle bending, respectively, derived from the force matching procedure for the cisplatin-DNA-*HMGB1* complex. r_{eq} ($\mathring{A}$) and θ_{eq} (degree) are equilibrium bond lengths and bond angles, respectively. See Fig. 5.5 for the atom type convention.

	Phase	No. of barriers	V_n
H7-N7-Pt-N5	180	3	0.073
H6-N6-Pt-N4	180	3	0.137
N7-Pt-N5-CK	0	4	2.340
N7-Pt-N5-CB	180	4	1.65
N6-Pt-N4-CK	0	4	3.02
N6-Pt-N4-CB	180	4	2.17

Table 5.5 Force constant of dihedral angles (kcal/mol) obtained from the force matching procedure. See Fig. 5.5 for the atom type convention.

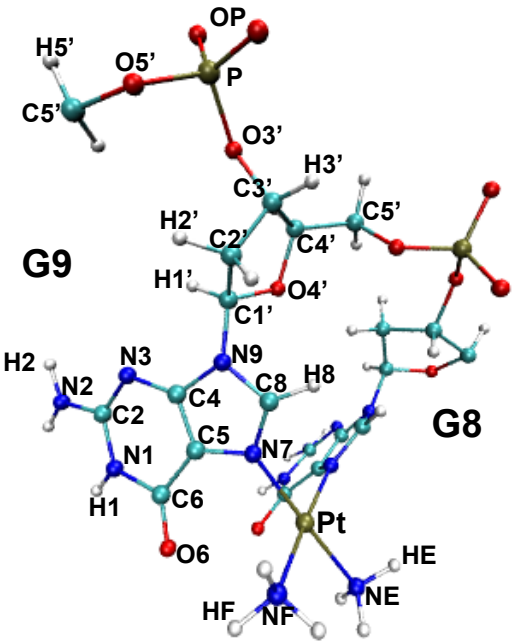


Figure 5.4 Atom name convention used for parameterization of partial charges.

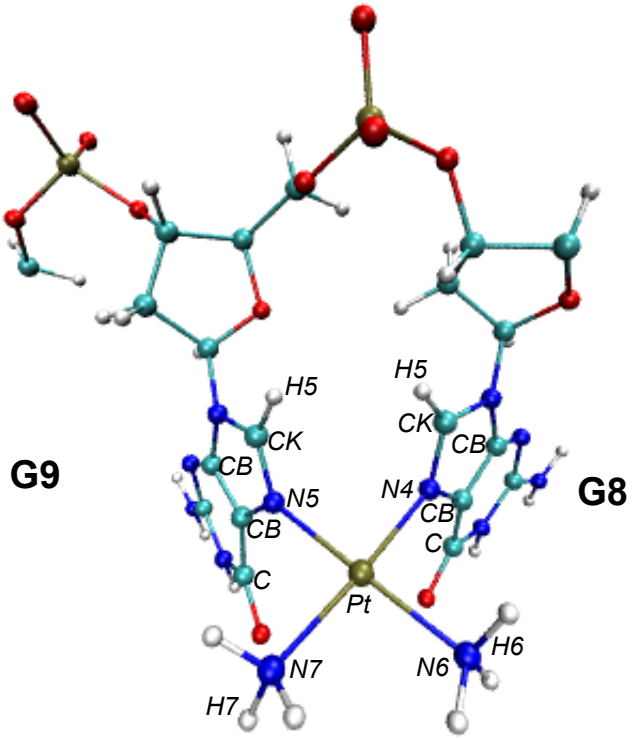


Figure 5.5 Atom type convention used for parameterization of bonded parameters.

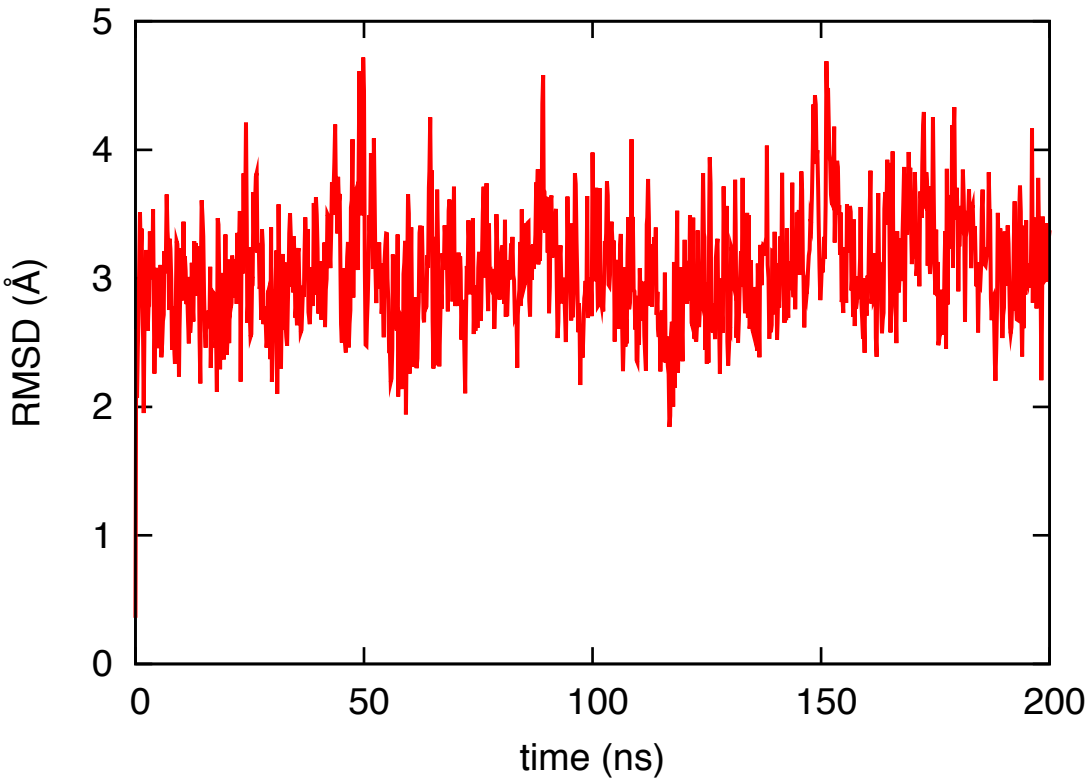


Figure 5.6 The backbone RMSD of classical MD trajectory.

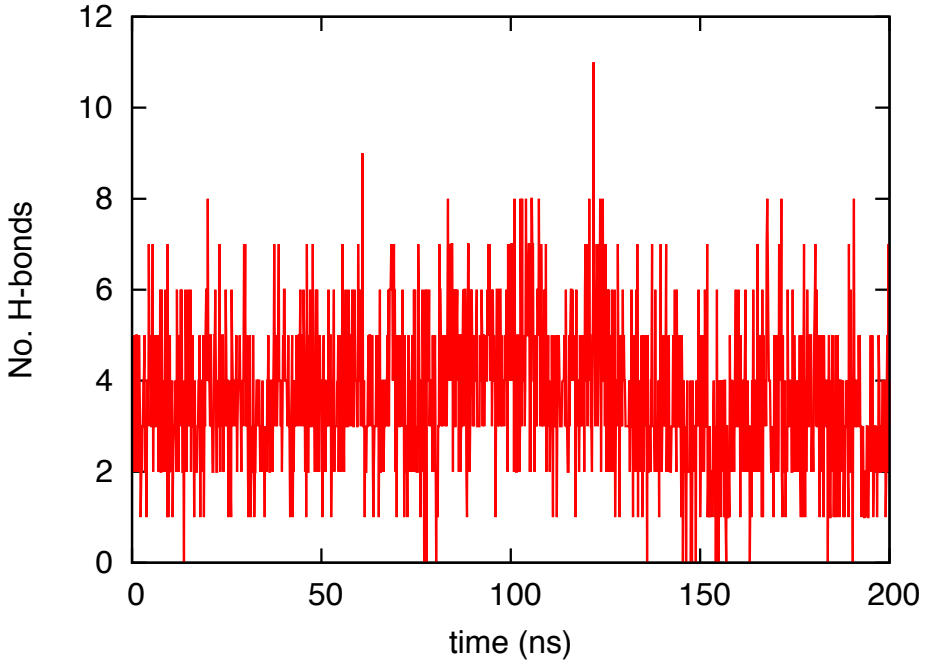


Figure 5.7 Number of hydrogen bonds between the platinated DNA and *HMGB1*

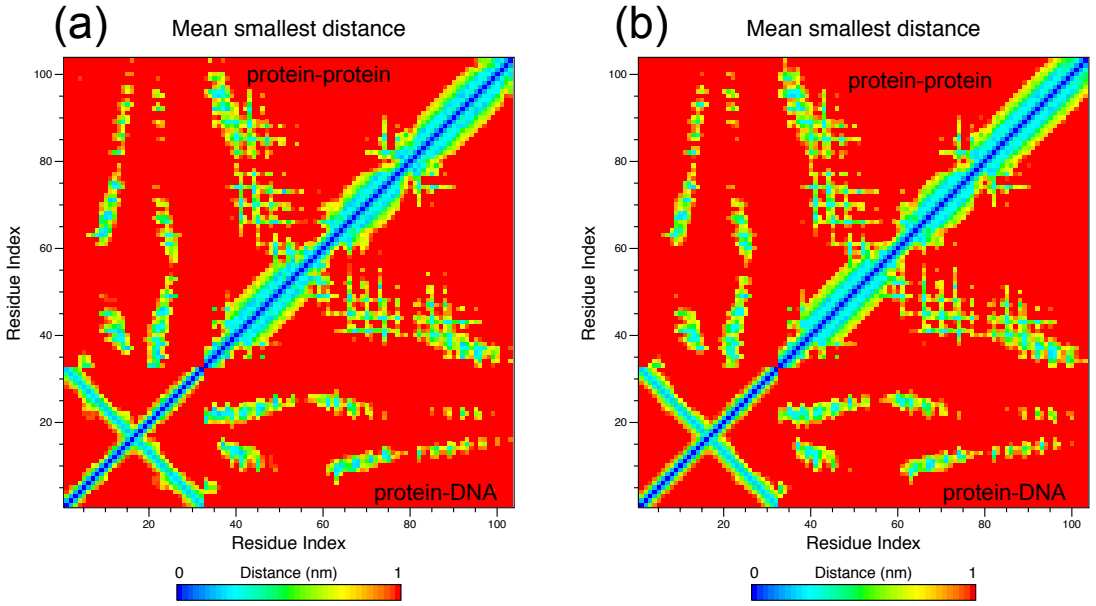


Figure 5.8 The contact map, which is defined as matrix of the smallest distances between pairs of residues, of the X-ray structure (a) and the classical MD trajectory (b).

6

Summary and outlooks

Cisplatin is one of the most potent drugs against testicular, ovarian, cervical, head and neck, esophageal, and small-cell lung cancers [1, 2, 4, 8]. Its beneficial effect arises from its binding to DNA [1, 2, 3, 4, 9]. Unfortunately, cisplatin suffers from several drawbacks, including drug resistance, which severely limits its efficacy, since only a small fraction of the drug reaches the target DNA [87, 30]. In addition, cisplatin-DNA adducts can be removed by Nucleotide Excision Repair (NER) proteins [2, 28, 86, 98]. Understanding and eventually overcoming these severe problems require a detailed knowledge of how cisplatin interacts with a variety of proteins in the cell. Whilst experiments are providing a great amount of knowledge on the molecular basis of resistance, computational biophysical studies on cisplatin resistance are basically lacking. This thesis has contributed to two aspects of cisplatin/protein interactions, which are relevant to the toxicity and the limited uptake of the drug, using computational physics tools.

(i) We have studied the binding of cisplatin and its isomer transplatin to a model peptide (named *Mets7*), that mimics one of the metal binding motifs found at the extracellular domain of *Ctr1*. *Mets7* has been shown to be an excellent model for investigating interactions of platinum drugs with *Ctr1 in vitro* and in-cell conditions. In order to characterize the structural and spectroscopic properties of the adducts of cisplatin/transplatin with *Mets7*, we have used an arsenal of computational tools, based on hybrid Car-Parrinello/molecular mechanics (QM/MM) simulations, force matching technique, replica exchange MD simulations and computational spectroscopy approaches. Our calculated spectroscopic properties such as NMR, EXAFS and CD have been compared with the corresponding experimental data to validate our predictions. This allows us to elucidate the different reaction modes of cisplatin and transplatin towards *Ctr1* and to propose the most relevant species. The molecular models arising from our calculations can serve as the molecular basis for further attempts to understand mechanism of cisplatin uptake mediated by *Ctr1*.

Other copper transport proteins such as the copper chaperone *Atox1* and copper pumps *ATP7A/ATP7B* have been found to bind cisplatin *in vitro* and *in vivo*. However, the structural information of these cisplatin-protein adducts is mostly lacking. Therefore, the computational strategies which have been employed here for structural prediction cisplatin/transplatin-*Mets7* adducts can be extended to these systems. Recently, a project focusing on the adducts of *Atox1* and *ATP7B* with cisplatin, based on the same computational protocol for *Ctr1*, has been approved by Deutsche Forschungsgemeinschaft (DFG).

(ii) We have carried out preliminary calculations of platinated DNA binding to the *HMGB1* protein. Resistance to cisplatin is also linked to the ability of the resistant cells to repair DNA damages induced by cisplatin via the NER repair pathway. Fortunately, *HMGB1* protein may be able to prevent NER by binding to the cisplatin-DNA adduct and shielding it from being repaired by NER enzymes. Understanding the binding mechanism of platinated DNA to *HMGB1* protein is, therefore, a fundamental step towards understanding molecular basis of cisplatin resistance mechanisms. Here we address this issue by investigating structural and energetic properties of cisplatin-DNA adduct binding to *HMGB1* protein using free energy calculations based on metadynamics. In this thesis, we have performed a QM/MM simulation and force matching procedure to parameterize the force field for the platinated moiety. Then the system has been equilibrated in aqueous solution for 200ns. These preliminary calculations are used to prepare the system for the more expensive free energy calculations.

From the preliminary results obtained in (ii) we plan to combine well-tempered metadynamics with parallel tempering to calculate the free energy profile of the binding of cisplatin-DNA adduct to *HMGB1* protein. The combined scheme provides an efficient way to explore very complex free energy landscapes and allows large free energy barrier to be easily overcome. Therefore, it allows studies of very complex processes such as molecular recognition, which requires extensive sampling of the association/disassociation processes. However, this approach is very computationally demanding. Therefore, we have requested computing resource from PRACE [293] and our proposal is under evaluation.

A

Appendix

A.1 DFT-based geometry optimization of the model compounds

The model compounds are $[Pt(SH_2)_3Cl]^+$, $[Pt(SH_2)_3H_2O]^{2+}$, $[Pt(SH_2)_3OH]^+$ and trans-$[Pt(SH_2)_2(NH_3)_2]^{2+}$ (Fig. A.1) which were energy-minimized at the DFT level using the CPMD program [188]. Exchange and correlation functional were those of Becke [195] and of Lee, Yang and Parr [196] (BLYP), respectively. Kohn-Sham orbitals were expanded in plane waves up to 70 Ry. The core electrons were described using norm-conserving pseudopotentials of the Martins-Troullier type [197]. Isolated system conditions have been applied [198]. GDIIS method [234] was used to optimize the structures of the model compounds with a convergence criterion of 10^{-6} for the largest component of the energy gradient. The results of structural parameters are reported in Tab. A.1.

A.2 Conversion of NMR chemical shifts

For ^{1}H, the NMR chemical shift was calculated as:

$$\delta^H(\mathrm{X}) = \sigma^H(\mathrm{TMS}) - \sigma^H(\mathrm{X}), \tag{A.1}$$

where $\sigma^H(\mathrm{TMS})$ is the ^{1}H shielding constant calculated for an isolated TMS molecule using the same DFT level [210]. Vibration and rotation can contribute significantly to the calculated ^{13}C NMR shielding constant of TMS [210]. It is, therefore, important to use the shielding constant of a simpler molecule (such as CH_4 or C_2H_6) as a computational reference. In this case, the chemical shielding constants were referenced indirectly to TMS via the calculated ^{13}C shielding constant and the experimental chemical shift of C_2H_6 [210]. Namely, this is achieved via the identity:

$$\delta^C(\mathrm{X}) = \sigma^C(\mathrm{TMS}) - \sigma^C(\mathrm{X}) = \delta^C_{\mathrm{exp}}(\mathrm{C_2H_6}) + \left[\sigma^C(\mathrm{C_2H_6}) - \sigma^C(\mathrm{X})\right]_{\mathrm{calc}}. \tag{A.2}$$

The value for $\delta^{C}_{\mathrm{exp}}(C_2H_6)$ was taken from ref. [238]. Due to the use of pseudopotentials to describe the interactions between valence and core electrons, the calculated ^{13}C chemical shielding constants do not account for the effect of the core electrons [239]. Such a constant is effectively canceled out by the equation A.2.

A.3 QM/MM simulations and ^{195}Pt NMR chemical shielding calculation of the reference compound $[PtCl_6]^{2-}$

The molecule $[PtCl_6]^{2-}$ was built as a new AMBER residue using Antechamber package [240]. The Gaussian 09 program [241] was used to optimize the structure of $[PtCl_6]^{2-}$ and to calculate the RESP charges [242]. The system was inserted in a box of 1000 water molecules. 2 Na^+ ions were added to neutralize the system. The 5 ns-long classical MD simulation was performed with the same setup as that of the preliminary simulations of the platinated peptides (see Sec. 4.2.1). During the MD simulation, the structure of $[PtCl_6]^{2-}$ was kept fixed. Then 10ps-long QM/MM simulation was carried out based on the last MD snapshot. QM part is the molecule and the MM part is the solvent and counterions. The QM/MM simulation procedure was the same as that of the platinated peptides (see Sec. 4.2.2). 50 equidistant snapshots were extracted from the QM/MM simulation for the ^{195}Pt chemical shielding calculation. The ^{195}Pt chemical shielding calculation procedure was the same as that for the platinated peptides (See Sec. 4.2.5.2).

A.4 Sampling of the QM region by QM/MM simulations

As shown in (Fig. A.3), the RMSD of the QM parts show that the local geometry around the platinum ion is stable during the QM/MM simulations. The RMSD of the T1 model shows a small increase near the end (at ∼20 ps) of the 25 ps-long trajectory (Fig. A.3). We compared the structures at the beginning of the simulation, after few ps and at the end of the simulation. The increase in RMSD is basically due to the change in the orientation of the two ammine groups. Indeed, the two ammine groups rotate around the S(Met1)-S(Met4) axis by approximately 15°. The rest of the T1 model essentially fluctuates around an average position. This movement does not affect the calculated values of NMR shifts reported in Tab. A.9 and A.11.

The sampling of the QM region was further checked by performing the principal component analysis and calculating the cosine contents of the first few principal components [224]. The cosine contents for all six systems are small and do not show any sign of poor sampling (Tab. A.7). It should be noted that the cosine content used here is a good indicator for bad sampling. If the cosine content has a large value (close to 1), then the sampling is surely not sufficient. However, the reverse is not necessarily true. A small value of the cosine content is a necessary but not sufficient condition for a good sampling.

A.5 Metadynamics method

The free energy as a function of d collective variables (Cvs)

$$S\left(\left\{\vec{R}_I\right\}\right) = \left\{S_1\left(\left\{\vec{R}_I\right\}\right), \ldots, S_d\left(\left\{\vec{R}_I\right\}\right)\right\} \tag{A.3}$$

is given by [284, 285]

$$F\left(S\right) = -k_B T \ln \int d\vec{R} d\vec{P} \exp\left(-\frac{H\left(\vec{R},\vec{P}\right)}{k_B T}\right) \delta\left(S - S\left(\vec{R}\right)\right). \tag{A.4}$$

The CVs should be chosen so as to distinguish between the initial and final state, describe all the relevant intermediates and include the slow modes, which cannot be satisfactorily sampled in the time scale of a standard MD simulation. There are several methods for evaluating $F(S)$. Among these, metadynamics adds an external history-dependent bias potential function of the CVs S and time t to the Hamiltonian:

$$H_{\mathrm{MetaD}} = H\left(\vec{R},\vec{P}\right) + V\left(S,t\right). \tag{A.5}$$

$V\left(S,t\right)$ is written as a sum of Gaussians of widths $\sigma_i (i = 1, \ldots, d)$ and height ω deposited along the system trajectory in the $S\left(\left\{\vec{R}_I\right\}\right)$ space to discourage the system from revisiting configurations that have already been sampled. $V\left(S,t\right)$ may be applied continuously during a MD simulation on the coordinates of the system:

$$V(S,t) = \omega \int_0^t dt' \exp\left\{-\frac{1}{2}\sum_{i=1}^{d}\left[\frac{S_i\left(\{\vec{R}_I\}\right) - S_i\left(\{\vec{R}_I(t')\}\right)}{\sigma_i}\right]^2\right\}. \tag{A.6}$$

Hence, metadynamics pushes the system away from local free energy minima, allowing it to explore reaction pathways since the system tends to escape the minima passing through the lowest free-energy saddle point.

Eventually, $V\left(S,t\right)$ provides the underlying free energy:

$$F(S) = -V\left(S, t \to \infty\right) + C, \tag{A.7}$$

where C is an irrelevant additive constant. Hence, metadynamics does not require a priori knowledge of the landscape. However, $-V(S,t)$ does not converge modulo a constant to the free energy, but oscillates around it. Hence, it is not trivial to decide when to stop a simulation. A solution to this problem has been provided by the *well-tempered* metadynamics [286]. Here the bias deposition rate decreases over simulation time, thus the dynamics becomes closer and closer to thermodynamic equilibrium as the simulation proceeds. The bias potential is given by

$$V\left(S,t\right) = k_B \Delta T \ln\left(1 + \frac{W}{\tau}\frac{N(S,t)}{k_B \Delta T}\right), \tag{A.8}$$

where $N(S,t)$ is the histogram of the S variables collected during the simulation, ΔT and W are two input parameters with the dimension of a temperature and an

energy, respectively, and τ is he time interval between two consecutive Gaussian depositions. Here, the bias potential converges in a single run. The probability distribution of the CVs can be shown to be proportional to:

$$P\left(S, t \to \infty\right) \propto \exp\left[-\frac{F(S)}{k_B\left(T+\Delta T\right)}\right]. \tag{A.9}$$

For $\Delta T \to 0$ and $\Delta T \to \infty$ we have standard MD and standard metadynamics, respectively. The free energy in well-tempered metadynamics can be shown to converge to:

$$F(S) = -\frac{T+\Delta T}{\Delta T} V\left(S, t \to \infty\right) + C. \tag{A.10}$$

To decrease the effect of neglecting slow degrees of freedom in the choice of the CVs, one can use a combination of metadynamics with parallel tempering (PT). In this combined scheme, M copies (replicas) of the same system undergo metadynamics at different temperatures $T_1 < T_2 < \ldots < T_M$. The exchange of configurations between two adjacent replicas (at T_i and T_{i+1}) is attempted with the probability

$$p\left(i \to i+1\right) = \min\left\{1, \Delta_{i,i+1}\right\}; i = 2, \ldots, M, \tag{A.11}$$

where

$$\begin{aligned}\Delta_{i,i+1} = \exp\Bigg[&\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_{i+1}}\right)\left(U\left(\left\{\vec{R}_I\right\}^i\right) - U\left(\left\{\vec{R}_I\right\}^{i+1}\right)\right) \\ &+ \frac{V^i\left(S\left(\left\{\vec{R}_I\right\}^i\right)\right) - V^i\left(S\left(\left\{\vec{R}_I\right\}^{i+1}\right)\right)}{k_B T_i} \\ &+ \frac{V^{i+1}\left(S\left(\left\{\vec{R}_I\right\}^i\right)\right) - V^{i+1}\left(S\left(\left\{\vec{R}_I\right\}^{i+1}\right)\right)}{k_B T_{i+1}}\Bigg]\end{aligned} \tag{A.12}$$

A.6 Supporting tables and figures

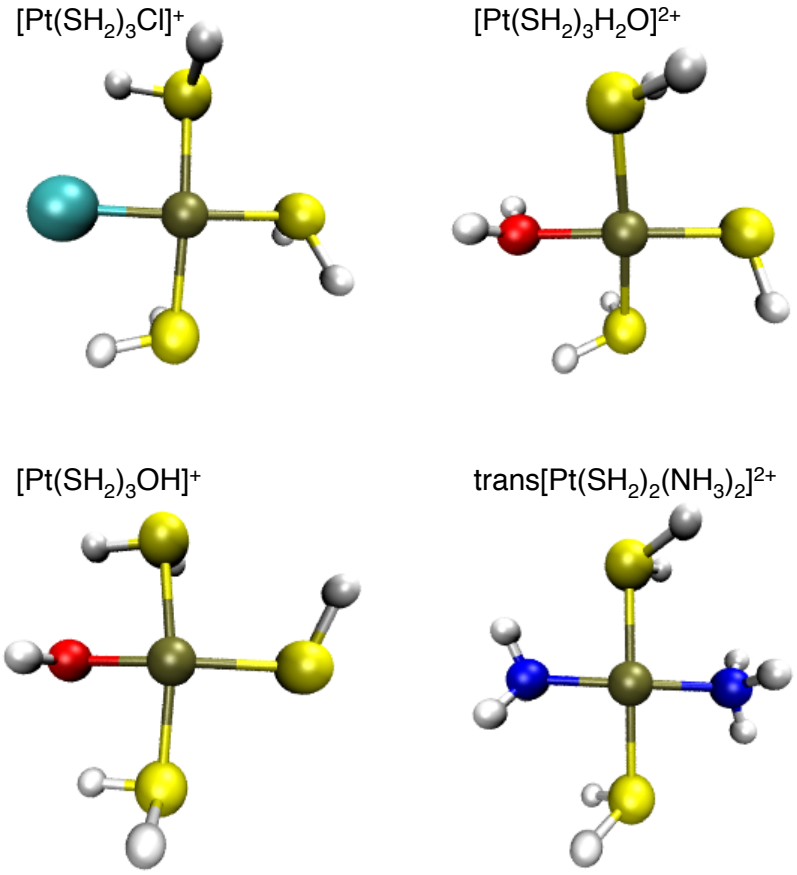


Figure A.1 DFT optimized model compounds

$[Pt(SH_2)_3Cl]^+$		$[Pt(SH_2)_3(H_2O)]^{2+}$		$[Pt(SH_2)_3(OH)]^+$		Trans-$[Pt(SH_2)_2(NH_3)_2]^{2+}$	
$Pt\text{-}S_{cis}$	2.39 Å	$Pt\text{-}S_{cis}$	2.42 Å	$Pt\text{-}S_{cis}$	2.38 Å	Pt-S	2.42 Å
$Pt\text{-}S_{trans}$	2.42 Å	$Pt\text{-}S_{trans}$	2.35 Å	$Pt\text{-}S_{trans}$	2.42 Å	Pt-N	2.11 Å
Pt-Cl	2.30 Å	Pt-O	2.15 Å	Pt-O	2.00 Å		
		O-H	0.99 Å	O-H	0.99 Å		
S_{cis}-Pt-S_{trans}	92°	S_{cis}-Pt-S_{trans}	92.0°	S_{cis}-Pt-S_{trans}	94.0°	S-Pt-N	90.0°
S_{cis}-Pt-Cl	88°	S_{cis}-Pt-O	88.0°	S_{cis}-Pt-O	86.0°		

Table A.1 Structural parameters of the model compounds optimized at the DFT level in gas phase. S_{cis} and S_{trans} are sulphur atoms cis and trans to the fourth ligand (chloride, water molecule, or OH^- group).

Atom names	C1			C2			C3		
Pt	0.3600			0.4075			0.4112		
Cl	-0.2777								
O				-0.4287			-0.7025		
H1/2				-0.1156			-0.1123		
Platinated Methionines									
	C1			C2			C3		
	Met1	Met4	Met7	Met1	Met4	Met7	Met1	Met4	Met7
CB	-0.1517	-0.1757	-0.2757	-0.2199	-0.1539	-0.2453	-0.1451	-0.0378	-0.2159
HB1/2	-0.0035	0.1004	0.4079	0.0137	0.0126	0.2371	0.1396	0.1762	0.0099
CG	-0.0003	0.0008	-0.1528	-0.0326	0.0423	-0.2000	0.0069	-0.0712	-0.0310
HG1/2	0.1384	0.1337	0.0006	0.0927	0.2328	-0.1800	0.0993	-0.0306	0.0819
SD	-0.2181	-0.2375	-0.4309	-0.1288	-0.2194	-0.1761	-0.1367	-0.3665	-0.1051
CE	0.0725	0.0260	-0.0701	0.2109	-0.0098	0.1619	0.0562	-0.0583	0.1303
HE1/2/3	0.2089	0.1031	0.0875	0.3872	0.1264	0.3452	0.1820	0.0870	0.2633

Table A.2 Charges (in fraction of electronic charge) derived from the force matching procedure for the cisplatin-*Mets7* adduct models. See Fig. A.2 for the atom name convention.

Atom names	T1		T2		T3	
Pt	0.4447		0.4068		0.4354	
N	-0.4228		-0.4382		-0.4051	
H	0.0158		0.0257		0.0315	
Platinated Methionines						
	T1		T2		T3	
	Met1	Met4	Met1	Met7	Met4	Met7
CB	-0.2427	-0.2965	-0.1805	-0.3827	-0.2229	-0.2229
HB2/3	0.3964	0.2336	0.0866	0.1761	0.2443	0.2443
CG	-0.1205	0.0390	0.0635	-0.0896	-0.0544	-0.0544
HG2/3	-0.0826	0.1916	0.2651	0.0761	-0.0015	-0.0015
SD	-0.2225	-0.0418	-0.2959	-0.1532	0.0991	0.0991
CE	0.0963	0.1334	0.0867	0.2167	0.1237	0.1237
HE1/2/3	0.2680	0.2667	0.2562	0.3983	0.2372	0.2372

Table A.3 Charges (in fraction of electronic charge) derived from the force matching procedure for the transplatin-*Mets7* adduct models. See Fig. A.2 for the atom name convention.

Cisplatin/Mets7								
	C1		C2		C3		AMBER	
	K_r	r_{eq}	K_r	r_{eq}			K_r	r_{eq}
Pt-Cl	105	2.343						
Pt-OE			82	2.131	127	2.044		
OE-HE			340	0.993	400	0.9844	553	0.957
Pt-SE	59	2.401	69	2.401	60	2.394		
Pt-ST	71	2.388	94	2.338	62	2.393		
CE-SE	102	1.901	106	1.890	98	1.898	227	1.81
CE-ST	99	1.895	94	1.897	93	1.895	227	1.81
CE-CE	198	1.540	205	1.543	194	1.540	310	1.526
Transplatin/Mets7								
	T1		T2		T3		AMBER	
	K_r	r_{eq}	K_r	r_{eq}	K_r	r_{eq}	K_r	r_{eq}
Pt-NE	107	2.093	108	2.095	106	2.096		
NE-HE	419	1.032	436	1.033	411	1.034	434	1.010
Pt-SE	67	2.394	75	2.379	69	2.388		
CE-SE	108	1.891	104	1.893	96	1.902	227	1.810
CE-CE	192	1.544	195	1.546	205	1.543	310	1.526

Table A.4 The harmonic force constant K_r of bond stretching (kcal/mol Å^{-2}) and equilibrium bond lengths r_{eq} (Å) obtained from the force matching procedure. See Fig. A.2 for the atom type convention.

Cisplatin/Mets7								
	C1		C2		C3		AMBER	
	K_θ	θ_{eq}	K_θ	θ_{eq}	K_θ	θ_{eq}	K_θ	θ_{eq}
CE-CE-SE	49	115.6	47	117.0	49	112.3	50	114.7
CE-SE-CE	68	101.0	56	105.1	61	102.0	62	98.9
CE-SE-Pt	47	109.0	36	109.5	42	108.2		
SE-Pt-ST	40	92.5	42	88.9	34	92.8		
SE-Pt-SE	21	175.0	30	177.8	26	180.0		
SE-Pt-Cl	14	87.5						
SE-Pt-OE			13	91.1	11	87.2		
CE-CE-ST	52	116.8	39	114.7	53	111.0	50	114.7
CE-ST-CE	57	102.0	48	97.4	66	101.7	62	98.9
CE-ST-Pt	40	112.3	46	113.8	43	110.9		
ST-Pt-Cl	31	180.0						
ST-Pt-OE			25	180.0	22	180.0		
Pt-OE-HE			19	118.4	26	106.7		
HE-OE-HE			36	107.3				
Transplatin/Mets7								
	T1		T2		T3		AMBER	
	K_θ	θ_{eq}	K_θ	θ_{eq}	K_θ	θ_{eq}	K_θ	θ_{eq}
CE-CE-SE	47	113.6	55	113.0	51	113.0	50	114.7
CE-SE-CE	66	104.4	64	103	54	105.5	62	98.9
CE-SE-Pt	42	109.2	34	111.3	37	107.4		
SE-Pt-SE	24	180.0	28	180.0	27	180.0		
SE-Pt-NE	46	90.0	38	90.0	39	90.0		
Pt-NE-HE	25	112.0	20	112.3	21	112.5		
HE-NE-HE	33	106.7	34	106.3	33	106.0	35	109.5
NE-Pt-NE	28	180.0	29	180.0	43	180.0		

Table A.5 The harmonic force constant of angle bending (kcal/mol rad^{-2}) and equilibrium bond angles (degree) obtained from the force matching procedure. See Fig. A.2 for the atom type convention.

Cisplatin/Mets7					
	Phase	No. of barriers	C1	C2	C3
Dihedral					
CE-SE-Pt-ST	180	4	1.697	3.041	0.706
SE-Pt-ST-CE	180	4	1.697	3.041	0.706
SE-Pt-OE-HE	0	4		0.036	0.000
Improper dihedral					
ST-S-Pt-Cl	180	2	4.792		
ST-S-Pt-OE	180	2		8.209	6.802
S-ST-Pt-S	180	2	7.520	2.793	3.634
Transplatin/Mets7					
	phase	No. of barriers	T1	T2	T3
Dihedral					
CE-SE-Pt-NE	0	4	2.693	1.347	3.623
SE-Pt-NE-HE	180	3	0.572	0.159	0.124
Improper dihedral					
S-NE-Pt-NE	180	2	10.543	3.793	3.902
S-NE-Pt-S	180	2	4.630	5.204	6.205

Table A.6 Force constant of dihedral angles (kcal/mol) obtained from the force matching procedure. See Fig. A.2 for the atom type convention.

	C1	C2	C3	T1	T2	T3
PC1	0.017	0.001	0.009	0.272	0.161	0.283
PC2	0.005	0.031	0.197	0.013	0.268	0.161
PC3	0.008	0.029	0.002	0.001	0.053	0.048
PC4	0.020	0.037	0.002	0.001	0.066	0.079

Table A.7 The cosine content of the first four principal components of the QM region in the QM/MM simulations.

^{13}C NMR				
		C1	C2	C3
RMSD w.r.t. the last REST snapshot (Å)		2.775	2.464	2.080
Met#	Atom Name			
Met1	C_γ	32.8 (5.6)	34.4 (4.1)	36.9 (4.4)
	C_ε	20.4 (6.4)	16.8 (3.9)	21.8 (4.4)
Met4	C_γ	38.7 (5.0)	37.9 (3.8)	40.1 (4.7)
	C_ε	16.9 (4.1)	25.3 (3.8)	19.4 (3.8)
Met7	C_γ	38.1 (6.7)	35.3 (3.6)	31.4 (4.2)
	C_ε	21.1 (4.6)	15.5 (2.4)	15.6 (4.1)
^{1}H NMR				
Met1	H_γ	2.92 (0.53)	2.99 (0.21)	3.07 (0.54)
	H_ε	2.90 (0.62)	2.58 (0.41)	2.93 (0.84)
Met4	H_γ	3.08 (0.58)	2.97 (0.32)	2.92 (0.66)
	H_ε	2.52 (0.69)	2.71 (0.34)	2.38 (0.57)
Met7	H_γ	3.68 (0.62)	3.59 (0.53)	3.13 (0.46)
	H_ε	2.70 (0.54)	2.41 (0.35)	2.74 (0.65)

Table A.8 NMR chemical shifts of C1-C3 (^{13}C (top) and ^{1}H (bottom)) obtained from QM/MM simulations before REST simulations.

^{13}C NMR				
		T1	T2	T3
RMSD w.r.t. the last REST snapshot (Å)		3.049	2.901	2.763
Met#	Atom Name			
Met1	C_γ	35.4 (4.9)	41.0 (5.3)	
	C_ε	20.4 (3.0)	24.8 (4.7)	
Met4	C_γ	33.0 (3.7)		35.5 (5.4)
	C_ε	17.0 (3.1)		18.0 (4.7)
Met7	C_γ		37.7 (4.7)	38.4 (6.2)
	C_ε		17.0 (4.2)	23.4 (4.4)
^{1}H NMR				
Met1	H_γ	3.34 (0.28)	3.25 (0.31)	
	H_ε	2.54 (0.34)	3.23 (0.54)	
Met4	H_γ	3.33 (0.28)		3.49 (0.38)
	H_ε	2.57 (0.34)		2.74 (0.40)
Met7	H_γ		4.06 (0.62)	3.49 (0.43)
	H_ε		2.58 (0.36)	2.81 (0.45)

Table A.9 NMR chemical shifts of T1-T3 (^{13}C (top) and ^{1}H (bottom)) obtained from QM/MM simulations before REST simulations.

^{13}C NMR		
Met#	Atom Name	
Met1	C_γ	30.7 (3.1)
	C_ε	17.2 (2.0)
Met4	C_γ	31.5 (2.5)
	C_ε	16.3 (2.7)
Met7	C_γ	31.2 (3.2)
	C_ε	15.1 (3.0)
^{1}H NMR		
Met1	H_γ	2.86 (0.40)
	H_ε	2.36 (0.31)
Met4	H_γ	2.91 (0.35)
	H_ε	2.33 (0.33)
Met7	H_γ	2.89 (0.38)
	H_ε	2.47 (0.39)

Table A.10 NMR chemical shifts of Apo (^{13}C (top) and ^{1}H (bottom)) obtained from QM/MM simulations for the representative snapshot of the REST trajectory. The RMSD of this representative snapshot relative to the last REST one is 3.3 Å.

Cisplatin-Mets7		Transplatin-Mets7	
Experiment	-4100	Experiment	-2900
C1	- 3750 (298)	T1	- 2879 (391)
C2	- 2734 (399)	T2	- 2785 (311)
C3	- 3460 (360)	T3	- 3066 (329)

Table A.11 ^{195}Pt NMR chemical shifts obtained from QM/MM simulations before REST simulations.

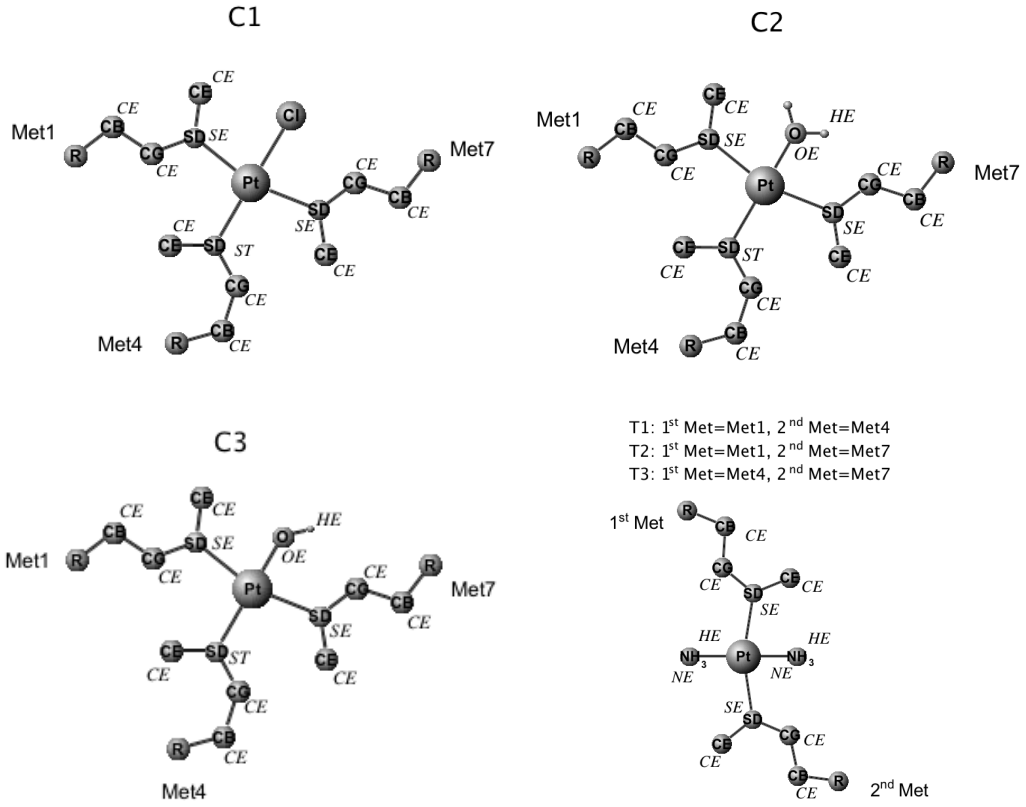


Figure A.2 Atom name and atom type convention for the QM atoms. The atom names are in bold and atom types in italics. Hydrogen atoms are not shown for the sake of clarity

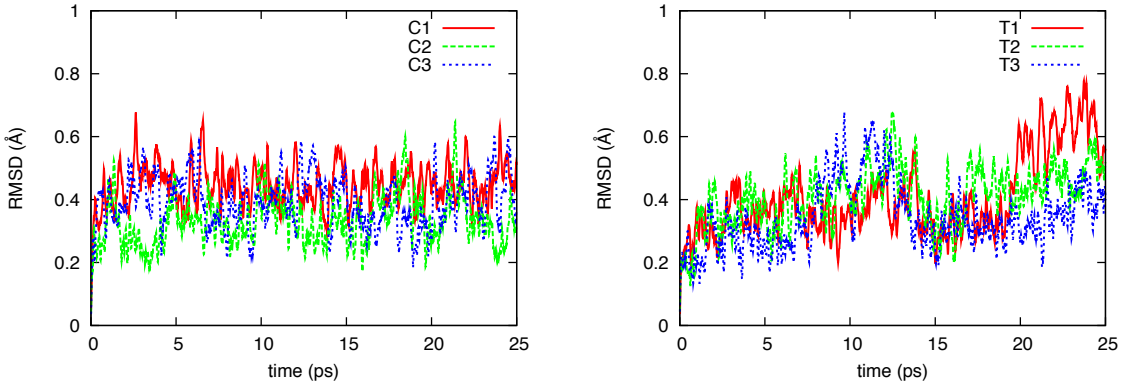


Figure A.3 RMSD of the QM region during the Car-Parrinello QM/MM simulations. Left: cisplatin-*Mets7*, right: transplatin-*Mets7*.

Bibliography

[1] Jung, Y.; Lippard, S. J. Direct cellular responses to platinum-induced DNA damage. *Chem Rev.* **2007**, *107*, 1387–1407.

[2] Kelland, L. The resurgence of platinum-based cancer chemotherapy. *Nat. Rev. Cancer* **2007**, *7*, 573–584.

[3] Wang, D.; Lippard, S. J. Cellular processing of platinum anticancer drugs. *Nat Rev Drug Discov.* **2005**, *4*, 307–320.

[4] Fuertes, M. A.; Alonso, C.; Perez, J. M. Biochemical modulation of Cisplatin mechanisms of action: enhancement of antitumor activity and circumvention of drug resistance. *Chem. Rev.* **2003**, *103*, 645–662.

[5] Rosenberg, B.; van Camp, L.; Krigas, T. INHIBITION OF CELL DIVISION IN ESCHERICHIA COLI BY ELECTROLYSIS PRODUCTS FROM A PLATINUM ELECTRODE. *Nature*, **1965**, *205*, 698–699.

[6] Rosenberg, B.; van Camp, L.; Trosko, J. E.; Mansour, V. H. Platinum compounds: a new class of potent antitumour agents. *Nature* **1969**, *222*, 385–386.

[7] Bosl, G. J.; Bajorin, D. F.; Sheinfeld, J.; Motzer, R. J.; Chaganti, R. S. K.; in Cancer: Principles & Practice of Oncology, ed. V. T. DeVita, V. T. Jr.; Hellman, S.; Rosenberg S. A., Lippincott Williams & Wilkins, Philadelphia, 6th ed, **2001**, pp. 1491–1518.

[8] Lippert, B. (Ed.) Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug, Wiley-VCH, Verlag Helvetica Chimica Acta, Postfach, CH8042 Zurich, Switzerland, **1999**.

[9] Jamieson, E. R.; Lippard, S. J. Structure, Recognition, and Processing of Cisplatin-DNA Adducts. *Chem. Rev.* **1999**, *99*, 2467–2498

[10] Pascoe, J. M.; Roberts, J. J. Interactions between mammalian cell DNA and inorganic platinum compounds. II. Interstrand cross-linking of isolated and cellular DNA by platinum(IV) compounds. *Biochem. Pharmacol.* **1974**, *23*, 1345–1357.

[11] Boulikas, T.; Vougiouka, M. Recent clinical trials using cisplatin, carboplatin and their combination chemotherapy drugs (review). *Oncol. Rep.* **2004**, *11*, 559–595.

[12] Bosl, G. J.; Motzer, R. J. Testicular germ-cell cancer. *N. Engl. J. Med.* **1997**, *337*, 242–253.

[13] Pil, P.; Lippard, S. J. Encyclopedia of Cancer, Vol. 1 Ed.: Bertino, J. R., Academic Press: San Diego, CA, **1997**.

[14] Wheate, N. J.; Walker, S.; Craig, G. E.; Oun, R. The status of platinum anticancer drugs in the clinic and in clinical trials. *Dalton Trans.* **2010**, *39*, 8113–8127.

[15] Reslova, S. The induction of lysogenic strains of Escherichia coli by cis-dichlorodiammineplatinum (II). *Chem. Biol. Interact.* **1971**, *4*, 66–70.

[16] Brouwer, J.; Van De Putte, P.; Fichtinger-Schepman, A. M. J.; Reedijk, J. Base-pair substitution hotspots in GAG and GCG nucleotide sequences in Escherichia coli K-12 induced by cis-diamminedichloroplatinum (II). *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 7010–7014.

[17] Popoff, S. C.; Beck, D. J.; Rupp, W. D. Repair of plasmid DNA damaged in vitro with cis- or trans-diamminedichloroplatinum(II) in Escherichia coli. *Mutat. Res.* **1987**, *183*, 129–137.

[18] Boffetta, P.; Fichtinger-Schepman, A. M. J.; Weiderpass, E.; van Dijk-Knijnenburg, H. C. M.; Stoter, G.; van Oosterom, A. T.; Keizer, H. J.; Fossa, S. D.; Kaldor, J.; Roy, P. Cisplatin-DNA adducts and protein-bound platinum in blood of testicular cancer patients. *Anticancer Drugs* **1998**, *9*, 125–129.

[19] Welters, M. J. P.; Fichtinger-Schepman, A. M. J.; Baan, R. A.; Jacobs-Bergmans, A. J.; Kegel, A.; van der Vijgh, W. J. F.; Braakhuis, B. J. M. Pharmacodynamics of cisplatin in human head and neck cancer: correlation between platinum content, DNA adduct levels and drug sensitivity in vitro and in vivo. *Br. J. Cancer* **1999**, *79*, 82–88.

[20] Cohen, G. L.; Ledner, J. A.; Bauer, W. R.; Ushay, H. M.; Caravana, C.; Lippard, S. J. Sequence dependent binding of cis-dichlorodiammineplatinum(II) to DNA. *J. Am. Chem. Soc.* **1980**, *102*, 2487-2488.

[21] Shimizu, M.; Rosenberg, B. A similar action to UV-irradiation and a preferential inhibition of DNA synthesis in E. coli by antitumor platinum compounds. *J. Antibiot.* **1973**, *26*, 243–245.

[22] Harder, H. C.; Smith, R. G.; Leroy, A. F. Template primer inactivation by cis- and trans-dichlorodiammine platinum for human DNA polymerase alpha, beta, and Rauscher murine leukemia virus reverse transcriptase, as a mechanism of cytotoxicity. *Cancer Res.* **1976**, *36*, 3821–3829.

[23] Davies, M. S.; Berners-Price, S. J.; Hambley, T. W. Rates of platination of -AG- and -GA- containing double-stranded oligonucleotides: effect of chloride concentration. *J. Inorg. Biochem.* **2000**, *79*, 167–172.

[24] Todd, R. C.; Lippard, S. J. Inhibition of transcription by platinum antitumor compounds. *Metallomics* **2009**, *1*, 280–291.

[25] Mymryk, J. S.; Zaniewski, E.; Archer, T. K. Cisplatin inhibits chromatin remodeling, transcription factor binding, and transcription from the mouse mammary tumor virus promoter in vivo. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2076–2080.

[26] Corda, Y.; Job, C.; Anin, M. F.; Leng, M.; Job, D. Spectrum of DNA–platinum adduct recognition by prokaryotic and eukaryotic DNA-dependent RNA polymerases. *Biochemistry* **1993**, *32*, 8582–8588.

[27] Mello, J. A.; Lippard, S. J.; Essigmann, J. M. DNA adducts of cis-diamminedichloroplatinum(II) and its trans isomer inhibit RNA polymerase II differentially in vivo. *Biochemistry* **1995**, *34*, 14783–14791.

[28] Siddik, Z. H. Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene* **2003**, *22*, 7265–7279.

[29] Arnesano, F.; Natile, G. Mechanistic insight into the cellular uptake and processing of cisplatin 30 years after its approval by FDA. *Coord. Chem. Rev.* **2009**, *253*, 2070–2081.

[30] Hall, M. D.; Okabe, M.; Shen, D. W.; Liang, X. J.; Gottesman, M. M. The role of cellular accumulation in determining sensitivity to platinum-based chemotherapy. *Annu. Rev. Pharmacol. Toxicol.* **2008**, *48*, 495–535.

[31] Lippard, S. J.; Berg, J. M. Principles of Bioinorganic Chemistry, University Science Books: Mill Valley, CA, **1994**.

[32] Kasherman, Y.; Sturup, S.; Gibson, D. Is glutathione the major cellular target of cisplatin? A study of the interactions of cisplatin with cancer cell extracts. *Med. Chem.* **2009**, *52*, 4319–4328.

[33] Blair, B. G.; Larson, C. A.; Safaei, R.; Howell, S. B. Copper transporter 2 regulates the cellular accumulation and cytotoxicity of Cisplatin and Carboplatin. *Clin. Cancer Res.* **2009**, *15*, 4312–4321.

[34] Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C.; Losacco, M.; Natile, G. Probing the interaction of cisplatin with the human copper chaperone Atox1 by solution and in-cell NMR spectroscopy. *J. Am. Chem. Soc.* **2011**, *133*, 18361–18369.

[35] Caradonna, J. P.; Lippard, S. J.; Gait, M. J.; Singh, M. J. The antitumor drug cis-dichlorodiammineplatinum forms an intrastrand d(GpG) crosslink upon reaction with [d(ApGpGpCpCpT)]2. *Am. Chem. Soc.* **1982**, *104*, 5793–5795.

[36] Fichtinger-Schepman, A. M. J.; van der Veer, J. L.; den Hartog, J. H. J.; Lohman, P. H. M.; Reedijk, J. Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. *Biochemistry* **1985**, *24*, 707–713.

[37] Eastman A. Reevaluation of interaction of cis-dichloro(ethylenediamine)platinum(II) with DNA. *Biochemistry* **1986**, *25*, 3912–3915.

[38] Terheggen, P. M. A. B.; Floot, B. G. J.; Scherer, E.; Begg, A. C.; Fichtinger-Schepman, A. M. J.; den Engelse, L. Immunocytochemical detection of interaction products of cis-diamminedichloroplatinum(II) and cis-diammine(1,1-cyclobutanedicarboxylato)platinum(II) with DNA in rodent tissue sections. *Cancer Res.* **1987**, *47*, 6719–6725.

[39] Lepre, C. A.; Strothkamp, K. G.; Lippard, S. J. Synthesis and 1H NMR spectroscopic characterization of trans-[Pt(NH3)2[d(ApGpGpCpCpT)-N7-A(1),N7-G(3)]]. *Biochemistry* **1987**, *26*, 5651–5657.

[40] Eastman, A,; Jennerwein M. M.; Nagel, D. L. Characterization of bifunctional adducts produced in DNA by trans-diamminedichloroplatinum(II). *Chem. Biol. Interact.* **1988**, *67*, 71–80.

[41] Takahara, P. M.; Rosenzweig, A. C.; Frederick C. A.; Lippard, S. J. Characterization of bifunctional adducts produced in DNA by trans-diamminedichloroplatinum(II). *Nature* **1995**, *377*, 649–652.

[42] Takahara, P. M.; Frederick, C. A.; Lippard, S. J. Single d(GpG)/cis-diammineplatinum(II) adduct-induced inhibition of DNA polymerization. *J. Am. Chem. Soc.* **1996**, *118*, 12309–12321.

[43] Yang, D.; van Boom, S. S. G. E.; Reedijk, J.; van Boom, J. H.; Wang, A. H.-J. Structure and isomerization of an intrastrand cisplatin-cross-linked octamer DNA duplex by NMR analysis. *Biochemistry* **1995**, *34*, 12912–12920.

[44] Gelasco, A.; Lippard, S. J. NMR solution structure of a DNA dodecamer duplex containing a cis-diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry* **1998**, *37*, 9230–9239.

[45] Magistrato, A.; Ruggerone, P.; Spiegel, K.; Carloni, P.; Reedijk, J. Binding of novel azole-bridged dinuclear platinum(II) anticancer drugs to DNA: insights from hybrid QM/MM molecular dynamics simulations. *J. Phys. Chem. B.* **2006**, *110*, 3604–3613.

[46] Spiegel, K.; Magistrato, A.; Carloni, P.; Reedijk, J.; Klein, M. L. Azole-bridged diplatinum anticancer compounds. Modulating DNA flexibility to escape repair mechanism and avoid cross resistance. *J. Phys. Chem. B.* **2007**, *111*, 11873–11876.

[47] Spiegel, K.; Rothlisberger, U.; Carloni, P. Cisplatin Binding to DNA Oligomers from Hybrid Car-Parrinello/Molecular Dynamics Simulations. *J. Phys. Chem. B* **2004**, *108*, 2699–2707.

[48] Robertazzi, A.; Platts, J. A. A QM/MM study of cisplatin-DNA oligonucleotides: from simple models to realistic systems. *Chemistry* **2006**, *12*, 5747–5756.

[49] Burda, J. V.; Leszczynski, J. How Strong Can the Bend Be on a DNA Helix from Cisplatin? DFT and MP2 Quantum Chemical Calculations of Cisplatin-Bridged DNA Purine Bases. *Inorg. Chem.* **2003**, *42*, 7162–7172.

[50] Pavelka, M.; Burda, J. V. Pt-bridges in various single-strand and double-helix DNA sequences. DFT and MP2 study of the cisplatin coordination with guanine, adenine, and cytosine. *J. Mol. Model.* **2007**, *13*, 367–379.

[51] Robertazzi, A.; Platts, J. A. Hydrogen bonding and covalent effects in binding of cisplatin to purine bases: ab initio and atoms in molecules studies. *Inorg. Chem.* **2005**, *44*, 267–274.

[52] Gkionis, K.; Platts, J. A. QM/MM studies of cisplatin complexes with DNA dimer and octamer. *Comput. Theor. Chem.* **2012**, *993*, 60–65.

[53] Spiegel, K.; Magistrato, A. Modeling anticancer drug-DNA interactions via mixed QM/MM molecular dynamics simulations. *Org. Biomol. Chem.* **2006**, *4*, 2507–2517.

[54] Tai, H. C.; Brodbeck, R.; Kasparkova, J.; Farrer, N. J.; Brabec, V.; Sadler, P. J.; Deeth, R. J. 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(N3)2(OH)2(pyridine)2]. *Inorg. Chem.* **2012**, *51*, 6830-6841.

[55] Suchánková, T.; Kubí¿ek, K.; Ka¿párková, J.; Brabec, V.; Kozelka, J. Platinum-DNA interstrand crosslinks: molecular determinants of bending and unwinding of the double helix. *J. Inorg. Biochem.* **2012**, *108*, 69–79.

[56] Bhattacharyya, D.; Ramachandran, S.; Sharma, S.; Pathmasiri, W.; King, C. L.; Baskerville-Abraham, I.;Boysen, G.; Swenberg, J. A.; Campbell, S. L.; Dokholyan, N. V.; Chaney, S. G. Flanking bases influence the nature of DNA distortion by platinum 1,2-intrastrand (GG) cross-links. *PLoS One.* **2011**, *6*, e23582.

[57] Sharma, S.; Gong, P.; Temple, B.; Bhattacharyya, D.; Dokholyan, N. V.; Chaney, S. G. Molecular dynamic simulations of cisplatin- and oxaliplatin-d(GG) intrastand cross-links reveal differences in their conformational dynamics. *J. Mol. Biol.* **2007**, *373*, 1123–1140.

[58] Téletchéa, S.; Skauge, T.; Sletten, E.; Kozelka, J. Cisplatin adducts on a GGG sequence within a DNA duplex studied by NMR spectroscopy and molecular dynamics simulations. *Chemistry* **2009**, *15*, 12320–12337.

[59] Scheeff, E. D.; Briggs, J. M.; Howell, S. B. Molecular modeling of the intrastrand guanine-guanine DNA adducts produced by cisplatin and oxaliplatin. *Mol. Pharmacol.* **1999**, *56*, 633–643.

[60] Baik, M. H.; Friesner, R. A.; Lippard, S. J. Theoretical study on the stability of N-glycosyl bonds: why does N7-platination not promote depurination? *J. Am. Chem. Soc.* **2002**, *124*, 4495–4503.

[61] Matsui, T.; Shigeta, Y.; Hirao, K. Multiple proton-transfer reactions in DNA base pairs by coordination of pt complex. *J. Phys. Chem. B.* **2007**, *111*, 1176–1181.

[62] Monjardet-Bas, V.; Elizondo-Riojas, M. A.; Chottard, J. C.; Kozelka, J. A combined effect of molecular electrostatic potential and N7 accessibility explains sequence-dependent binding of cis-[Pt(NH3)2(H2O)2]2+ to DNA duplexes. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2998–3001.

[63] Nguyen, T. H.; Arnesano, F.; Scintilla, S.; Rossetti, G.; Ippoliti, E.; Carloni, P.; Natile, G. Structural Determinants of Cisplatin and Transplatin Binding to the Met-Rich Motif of Ctr1: A Computational Spectroscopy Approach. *J. Chem. Theory Comput.* **2012**, *8*, 2912–2920.

[64] Van Garderen, C. J.; Van Houte, L. P. A. The solution structure of a DNA duplex containing the cis-Pt(NH3)2[d(-GTG-)-N7(G),N7(G)] adduct, as determined with high-field NMR and molecular mechanics/dynamics. *Eur. J. Biochem.* **1994**, 225, 1169–1179.

[65] Teuben, J.-M.; Bauer, C.; Wang, A. H.-J.; Reedijk, J. Solution structure of a DNA duplex containing a cis-diammineplatinum(II) 1,3-d(GTG) intrastrand cross-link, a major adduct in cells treated with the anticancer drug carboplatin. *Biochemistry* **1999**, *38*, 12305–12312.

[66] Hughes, E. N.; Engelsberg, B. N.; Billings, P. C. Purification of nuclear proteins that bind to cisplatin-damaged DNA. Identity with high mobility group proteins 1 and 2. *J. Biol. Chem.* **1992**, *267*, 13520–13527.

[67] Pil, P. M.; Lippard, S. J. Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science* **1992**, *256*, 234–247.

[68] Baxevanis, A. D.; Landsman, D. The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res.* **1995**, *23*, 1604–1613.

[69] Thomas, J. O.; Travers, A. A. HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem. Sci.* **2001**, *26*, 167–174.

[70] Huang, J.-C.; Zamble, D. B.; Reardon, J. T.; Lippard, S. J.; Sancar, A. HMG-domain proteins specifically inhibit the repair of the major DNA adduct of the anticancer drug cisplatin by human excision nuclease. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10394–10398.

[71] Zamble, D. B.; Mu, D.; Reardon, J. T.; Sancar, A.; Lippard, S. J. Repair of cisplatin–DNA adducts by the mammalian excision nuclease. *Biochemistry* **1996**, *35*, 10004–10013.

[72] Ugrinova, I.; Zlateva, S.; Pashev, I. G.; Pasheva, E. A. Native HMGB1 protein inhibits repair of cisplatin-damaged nucleosomes in vitro. *Int. J. Biochem. Cell Biol.* **2009**, *41*, 1556–1562.

[73] Mitkova, E.; Ugrinova, I.; Pashev, I. G.; Pasheva, E. A. The inhibitory effect of HMGB-1 protein on the repair of cisplatin-damaged DNA is accomplished through the acidic domain. *Biochemistry* **2005**, *44*, 5893–5898.

[74] He, Q.; Liang, C. H.; Lippard, S. J. Steroid hormones induce HMG1 overexpression and sensitize breast cancer cells to cisplatin and carboplatin. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5768–5772.

[75] Arioka, H.; Nishio, K.; Ishida, T.; Fukumoto, H.; Fukuoka, K.; Nomoto, T.; Kurokawa, H.; Yokote, H.; Abe, S.; Saijo, N. Enhancement of cisplatin sensitivity in high mobility group 2 cDNA-transfected human lung cancer cells. *Jpn. J. Cancer Res.* **1999**, *90*, 108–115.

[76] Ohndorf, U. M.; Rould, M. A.; He, Q.; Pabo, C. O.; Lippard, S. J. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature* **1999**, *399*, 708–712.

[77] Hardman, C. H.; Broadhurst, R. W.; Raine, A. R. C.; Grasser, K. D.; Thomas, J. O.; Laue, E. D. Structure of the A-domain of HMG1 and its interaction with DNA as studied by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry* **1995**, *34*, 16596–16607.

[78] Park, S.; Lippard, S. J. Redox state-dependent interaction of HMGB1 and cisplatin-modified DNA. *Biochemistry* **2011**, *50*, 2567–2574.

[79] Perez, R. P. Cellular and molecular determinants of cisplatin resistance. *Eur. J. Cancer* **1998**, *34*, 1535–1542.

[80] Muggia, F. M.; Los, G. Platinum resistance: laboratory findings and clinical implications. *Stem Cells* **1993**, *11*, 182–193.

[81] Wernyj, R. P; Morin, P. J. Molecular mechanisms of platinum resistance: still searching for the Achilles' heel. *Drug Resist. Updates* **2004**, *7*, 227–232.

[82] Kuo, M. T.; Chen, H. H.; Song, I. S.; Savaraj, N.; Ishikawa, T. The roles of copper transporters in cisplatin resistance. *Cancer Metastasis Rev.* **2007**, *26*, 71–83.

[83] Gately, D. P.; Howell, S. B. Cellular accumulation of the anticancer agent cisplatin: a review. *Br. J. Cancer* **1993**, *67*, 1171–1176.

[84] Suzuki, T.; Nishio, K.; Tanabe, S. The MRP family and anticancer drug metabolism. *Curr. Drug Metab.* **2001**, *2*, 367–377.

[85] Stewart, D. J. Mechanisms of resistance to cisplatin and carboplatin. *Crit. Rev. Oncol. Hematol.* **2007**, *63*, 12–31.

[86] Kartalou, M.; Essigmann, J. M. Mechanisms of resistance to cisplatin. *Mutat Res.* **2001**, *478*, 23–43.

[87] Andrews, P. A.; Howell, S. B. Cellular pharmacology of cisplatin: perspectives on mechanisms of acquired resistance. *Cancer Cells* **1990**, *2*, 35–43.

[88] Johnson S. W.; Shen D-W.; Pastan, I.; Gottesman, M. M.; Hamilton, T. C. Cross-resistance, cisplatin accumulation, and platinum-DNA adduct formation and removal in cisplatin-sensitive and -resistant human hepatoma cell lines. *Exp. Cell Res.* **1996**, *226*, 133–139.

[89] Koga, H.; Kotoh, S.; Nakashima, M.; Yokomizo, A.; Tanaka, M.; Naito, S. Accumulation of intracellular platinum is correlated with intrinsic cisplatin resistance in human bladder cancer cell lines. *Int. J. Oncol.* **2000**, *16*, 1003–1007.

[90] Shen D-W.; Goldenberg, S.; Pastan, I.; Gottesman, M. M. Decreased accumulation of [14C]carboplatin in human cisplatin-resistant cells results from reduced energy-dependent uptake. *J. Cell Physiol.* **2000**, *183*, 108–116.

[91] Hector, S.; Bolanowska-Higdon, W.; Zdanowicz, J.; Hitt, S.; Pendyala, L. In vitro studies on the mechanisms of oxaliplatin resistance. *Cancer Chemother. Pharmacol.* **2001**, *48*, 398–406.

[92] Fuertes, M. A.; Castilla, J.; Alonso, C.; Pe rez, J. M. Novel concepts in the development of platinum antitumor drugs. *Curr. Med. Chem. Anticancer Agents* **2002**, *2*, 539–552.

[93] Ishikawa, T. The ATP-dependent glutathione S-conjugate export pump. *Trends Biochem. Sci.* **1992**, *17*, 463–468.

[94] Mistry, P.; Kelland, L. R.; Abel, G.; Sidhar, S.; Harrap, K. R. The relationships between glutathione, glutathione-S-transferase and cytotoxicity of platinum drugs and melphalan in eight human ovarian carcinoma cell lines. *Br. J. Cancer* **1991**, *64*, 215–220.

[95] Lewis, A. D.; Hayes, J. D.; Wolf, C. R. Glutathione and glutathione-dependent enzymes in ovarian adenocarcinoma cell lines derived from a patient before and after the onset of drug resistance: intrinsic differences and cell cycle effects. *Carcinogenesis* **1988**, *9*, 1283–1287.

[96] Kelley, S. L.; Basu, A.; Teicher, B. A.; Hacker, M. P.; Hamer, D. H.; Lazo, J. S. Overexpression of metallothionein confers resistance to anticancer drugs. *Science* **1988**, *241*, 1813–1815.

[97] Holford, J.; Beale, P. J.; Boxall, F. E.; Sharp, S. Y.; Kelland, L. R. Mechanisms of drug resistance to the platinum complex ZD0473 in ovarian cancer cell lines. *Eur. J. Cancer* **2000**, *36*, 1984–1990.

[98] Brabec, V.; Kasparkova, J. Modifications of DNA by platinum complexes. Relation to resistance of tumors to platinum antitumor drugs. *Drug Resist. Updat.* **2005**, *8*, 131–146.

[99] Johnson, S. W.; Swiggard, P. A.; Handel, L. M.; Brennan, J. M.; Godwin, A. K.; Ozols, R. F.; Hamilton, T. C. Relationship between platinum-DNA adduct formation and removal and cisplatin cytotoxicity in cisplatin-sensitive and -resistant human ovarian cancer cells. *Cancer Res.* **1994**, *54*, 5911–5916.

[100] Gonzalez, V. M.; Fuertes, M. A.; Alonso, C.; Perez, J. M. Is cisplatin-induced cell death always produced by apoptosis? *Mol. Pharmacol.* **2001**, *59*, 657–663.

[101] Reardon, J. T.; Vaisman, A.; Chaney, S. G.; Sancar, A. Efficient nucleotide excision repair of cisplatin, oxaliplatin, and Bis-aceto-ammine-dichloro-cyclohexylamine-platinum(IV) (JM216) platinum intrastrand DNA diadducts. *Cancer Res.* **1999**, *59*, 3968–3971.

[102] Chaney, S. G.; Sancar, A. DNA repair: enzymatic mechanisms and relevance to drug response. *J. Natl. Cancer. Inst.* **1996**, *88*, 1346–1360.

[103] Kelland, L. R.; Mistry, P.; Abel, G.; Friedlos, F.; Loh, S. Y.; Roberts, J. J.; Harrap, K. R. Establishment and characterization of an in vitro model of acquired resistance to cisplatin in a human testicular nonseminomatous germ cell line. *Cancer Res.* **1992**, *52*, 1710–1716.

[104] Zamble, D. B.; Mikata,Y.; Eng, C. H.; Sandman, K. E.; Lippard, S. J. Testis-specific HMG-domain protein alters the responses of cells to cisplatin. *J. Inorg. Biochem.* **2002**, *91*, 451–462.

[105] Andrews, P. A.; Velury, S.; Mann, S. C.; Howell, S. B. cis-Diamminedichloroplatinum(II) accumulation in sensitive and resistant human ovarian carcinoma cells. *Cancer Res.* **1988**, *48*, 68–73.

[106] Eichholtz-Wirth, H.; Hietel, B. The relationship between cisplatin sensitivity and drug uptake into mammalian cells in vitro. *Br. J. Cancer* **1986**, *54*, 239–243.

[107] Gale, G. R.; Morris, C. R.; Atkins, L. M.; Smith, A. B. Binding of an antitumor platinum compound to cells as influenced by physical factors and pharmacologically active agents. *Cancer Res.* **1973**, *33*, 813–818.

[108] Hecquet, B.; Leroy, A.; Lefebvre, J. L.; Peyrat, J. P.; Adenis, L. Uptake of platinum compounds in human tumors. In vitro study. *Bull. Cancer* **1986**, *73*, 535–541.

[109] Hromas, R. A.; North, J. A.; Burns, C. P. Decreased cisplatin uptake by resistant L1210 leukemia cells. *Cancer Lett.* **1987**, *36*, 197–201.

[110] Howell, S. B.; Safaei, R.; Larson, C. A.; Sailor, M. J. Copper transporters and the cellular pharmacology of the platinum-containing cancer drugs. *Mol Pharmacol.* **2010**, *77*, 887–894.

[111] Gupta, A.; Lutsenko, S. Human copper transporters: mechanism, role in human diseases and therapeutic potential. *Future Med Chem.* **2009**, *1*, 1125–1142.

[112] Prohaska, J. R.; Gybina, A. A. Intracellular copper transport in mammals. *J. Nutr.* **2004**, *134*, 1003–1006.

[113] Prohaska, J. R. Role of copper transporters in copper homeostasis. *Am. J. Clin. Nutr.* **2008**, *88*, 826S–829S.

[114] Sharp, P. A. Ctr1 and its role in body copper homeostasis. *Int. J. Biochem. Cell. Biol.* **2003**, *35*, 288–291.

[115] Wessling-Resnick, M. Understanding copper uptake at the molecular level. *Nutr. Rev.* **2002**, *60*, 177–179.

[116] Lin, X.; Okuda, T.; Holzer, A.; Howell, S. B. The copper transporter CTR1 regulates cisplatin uptake in Saccharomyces cerevisiae. *Mol. Pharmacol.* **2002**, *62*, 1154–1159.

[117] Holzer, A. K.; Katano, K.; Klomp, L. W.; Howell, S. B. Cisplatin rapidly down-regulates its own influx transporter hCTR1 in cultured human ovarian carcinoma cells. *Clin. Cancer Res.* **2004**, *10*, 6744–6749.

[118] Holzer, A. K.; Samimi, G.; Katano, K.; Naerdemann, W.; Lin, X.; Safaei, R.; Howell, S. B. The copper influx transporter human copper transport protein 1 regulates the uptake of cisplatin in human ovarian carcinoma cells. *Mol. Pharmacol.* **2004**, *66*, 817–823.

[119] Holzer, A. K.; Manorek, G. H.; Howell, S. B. Contribution of the major copper influx transporter CTR1 to the cellular accumulation of cisplatin, carboplatin, and oxaliplatin. *Mol. Pharmacol.* **2006**, *70*, 1390–1394.

[120] Holzer, A. K.; Varki, N. M.; Le, Q. T.; Gibson, M. A.; Naredi, P.; Howell, S. B. Expression of the human copper influx transporter 1 in normal and malignant human tissues. *J. Histochem. Cytochem.* **2006**, *54*, 1041–1049.

[121] Holzer, A. K.; Howell, S. B. The internalization and degradation of human copper transporter 1 following cisplatin exposure. *Cancer Res.* **2006**, *66*, 10944–10952.

[122] Larson, C. A.; Blair, B. G.; Safaei, R.; Howell, S. B. The role of the mammalian copper transporter 1 in the cellular accumulation of platinum-based drugs. *Mol. Pharmacol.* **2009**, *75*, 324–330.

[123] Noordhuis, P.; Laan, A. C.; van de Born, K.; Losekoot, N.; Kathmann, I.; Peters, G. J. Oxaliplatin activity in selected and unselected human ovarian and colorectal cancer cell lines. *Biochem. Pharmacol.* **2008**, *76*, 53–61.

[124] Liu, J. J.; Jamieson, S. M.; Subramaniam, J. Ip V.; Jong, N. N.; Mercer, J. F.; McKeage, M. J. Neuronal expression of copper transporter 1 in rat dorsal root ganglia: association with platinum neurotoxicity. *Cancer Chemother. Pharmacol.* **2009**, *64*, 847–856.

[125] Linz, R.; Lutsenko, S. Copper-transporting ATPases ATP7A and ATP7B: cousins, not twins. *J. Bioenerg. Biomembr.* **2007**, *39*, 403–407.

[126] Lutsenko, S.; Barnes, N. L.; Bartee, M. Y.; Dmitriev, O. Y. Function and regulation of human copper-transporting ATPases. *Physiol. Rev.* **2007**, *87*, 1011–1046.

[127] Luk, E.; Jensen, L. T.; Culotta, V. C. The many highways for intracellular trafficking of metals. *J. Biol. Inorg. Chem.* **2003**, *8*, 803–809.

[128] Safaei, R.; Holzer, A. K.; Katano, K.; Samimi, G.; Howell, S. B. The role of copper transporters in the development of resistance to Pt drugs. *J. Inorg. Biochem.* **2004**, *98*, 1607–1613.

[129] Katano, K.; Safaei, R.; Samimi, G.; Holzer, A.; Rochdi, M.; Howell, S. B. The copper export pump ATP7B modulates the cellular pharmacology of carboplatin in ovarian carcinoma cells. *Mol. Pharmacol.* **2003**, *64*, 466–473.

[130] Katano, K.; Safaei, R.; Samimi, G.; Holzer, A.; Tomioka, M.; Goodman, M.; Howell, S. B. Confocal microscopic analysis of the interaction between cisplatin and the copper transporter ATP7B in human ovarian carcinoma cells. *Clin. Cancer Res.* **2004**, *10*, 4578–4588.

[131] Safaei, R.; Howell, S. B. Copper transporters regulate the cellular pharmacology and sensitivity to Pt drugs. *Crit. Rev. Oncol. Hematol.* **2005**, *53*, 13–23.

[132] Safaei, R. Role of copper transporters in the uptake and efflux of platinum containing drugs. *Cancer Lett.* **2006**, *234*, 34–39.

[133] Samimi, G.; Katano, K.; Holzer, A. K.; Safaei, R.; Howell, S. B. Modulation of the cellular pharmacology of cisplatin and its analogs by the copper exporters ATP7A and ATP7B. *Mol. Pharmacol.* **2004**, *66*, 25–32.

[134] Safaei, R.; Katano, K.; Larson, B. J.; Samimi, G.; Holzer, A. K.; Naerdemann, W.; Tomioka, M.; Goodman, M.; Howell, S. B. Intracellular localization and trafficking of fluorescein-labeled cisplatin in human ovarian carcinoma cells. *Clin. Cancer Res.* **2005**, *11*, 756–767.

[135] Klomp, L. W.; Lin, S. J.; Yuan, D. S.; Klausner, R. D.; Culotta, V. C.; Gitlin, J. D. Identification and functional expression of HAH1, a novel human gene involved in copper homeostasis. *J. Biol. Chem.* **1997**, *272*, 9221–9226.

[136] Hamza, I.; Schaefer, M.; Klomp, L. W.; Gitlin, J. D. Interaction of the copper chaperone HAH1 with the Wilson disease protein is essential for copper homeostasis. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13363–13368.

[137] Hamza, I.; Prohaska, J.; Gitlin, J. D. Essential role for Atox1 in the copper-mediated intracellular trafficking of the Menkes ATPase. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1215–1220.

[138] Safaei, R.; Maktabi, M. H.; Blair, B. G.; Larson, C. A.; Howell, S. B. Effects of the loss of Atox1 on the cellular pharmacology of cisplatin. *J. Inorg. Biochem.* **2009**, *103*, 333–341.

[139] Itoh, S.; Kim, H. W.; Nakagawa, O.; Ozumi, K.; Lessner, S. M.; Aoki, H.; Akram, K.; McKinney, R. D.; Ushio-Fukai, M.; Fukai, T. Novel role of antioxidant-1 (Atox1) as a copper-dependent transcription factor involved in cell proliferation. *J. Biol. Chem.* **2008**, *283*, 9157–9167.

[140] Muller, P. A.; Klomp, L. W. ATOX1: a novel copper-responsive transcription factor in mammals? *Int. J. Biochem. Cell Biol.* **2009**, *41*, 1233–1236.

[141] Boal, A. K.; Rosenzweig, A.C. Crystal structures of cisplatin bound to a human copper chaperone. *J. Am. Chem. Soc.* **2009**, *131*, 14196–14197.

[142] Arnesano, F.; Scintilla, S.; Natile, G. Interaction between platinum complexes and a methionine motif found in copper transport proteins. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 9062–9064.

[143] Arnesano, F.; Natile, G. "Platinum on the road": Interactions of antitumoral cisplatin with proteins. *Pure Appl. Chem.* **2008**, *80*, 2715–2725.

[144] Jiang, J.; Nadas, I. A.; Kim, M. A.; Franz, K. J. A Mets motif peptide found in copper transport proteins selectively binds Cu(I) with methionine-only coordination. *Inorg. Chem.* **2005**, *44*, 9787–9787.

[145] Nose, Y.; Rees, E. M.; Thiele, D. J. Structure of the Ctr1 copper trans'PORE'ter reveals novel architecture. *Trends Biochem. Sci.* **2006**, *31*, 604–607.

[146] Wu, Z.; Liu, Q.; Liang, X.; Yang, X.; Wang, N.; Wang, X.; Sun, H.; Lu, Y.; Guo, Z. J. Reactivity of platinum-based antitumor drugs towards a Met- and His-rich 20mer peptide corresponding to the N-terminal domain of human copper transporter 1. *Biol. Inorg. Chem.* **2009**, *14*, 1313–1323.

[147] Crider, S. E.; Holbrook, R. J.; Franz, K. J. Coordination of platinum therapeutic agents to met-rich motifs of human copper transport protein1. *Metallomics* **2010**, *2*, 74–83.

[148] Banci, L.; Bertini, I.; Cantini, F.; Rosenzweig, A. C.; Yatsunyk, L. A. Metal binding domains 3 and 4 of the Wilson disease protein: solution structure and interaction with the copper(I) chaperone HAH1. *Biochemistry* **2008**, *47*, 7423–7429.

[149] Achila, D.; Banci, L.; Bertini, I.; Bunce, J.; Ciofi-Baffoni, S.; Huffman, D. L. Structure of human Wilson protein domains 5 and 6 and their interplay with domain 4 and the copper chaperone HAH1 in copper uptake. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5729–5734.

[150] Wernimont, A. K.; Huffman, D. L.; Lamb, A. L.; OHalloran, T. V.; Rosenzweig, A. C. Structural basis for copper transfer by the metallochaperone for the Menkes/Wilson disease proteins. *Nat. Struct. Biol.* **2000**, *7*, 766–771.

[151] Anastassopoulou, I.; Banci, L.; Bertini, I.; Cantini, F.; Katsari, E.; Rosato, A. Solution structure of the apo and copper(I)-loaded human metallochaperone HAH1. *Biochemistry* **2004**, *43*, 13046–13053.

[152] Dmitriev, O. Y. Mechanism of tumor resistance to cisplatin mediated by the copper transporter ATP7B. *Biochem Cell Biol.* **2011**, *89*, 138–147.

[153] Dolgova, N. V.; Olson, D.; Lutsenko, S.; Dmitriev, O. Y. The soluble metal-binding domain of the copper transporter ATP7B binds and detoxifies cisplatin. *Biochem J.* **2009**, *419*, 51–56.

[154] Leonhardt, K.; Gebhardt, R.; Mössner, J.; Lutsenko, S.; Huster, D. Functional interactions of Cu-ATPase ATP7B with cisplatin and the role of ATP7B in the resistance of cells to the drug. *J. Biol. Chem.* **2009**, *284*, 7793–7802.

[155] Lopes, J. F.; Rocha, W. R.; Dos Santos, H. F.; De Almeida, W. B. Theoretical study of the potential energy surface for the interaction of cisplatin and their aquated species with water. *J. Chem. Phys.* **2008**, *128*, 165103–165116.

[156] Sutter, K.; Truflandier, L. A.; Autschbach, J. NMR J-coupling constants in cisplatin derivatives studied by molecular dynamics and relativistic DFT. *ChemPhysChem* **2011**, *12*, 1448–1455.

[157] Truflandier, L. A.; Sutter, K.; Autschbach, J. Solvent effects and dynamic averaging of 195Pt NMR shielding in cisplatin derivatives. *Inorg. Chem.* **2011**, *50*, 1723–1732.

[158] Lau, J. K.; Ensing, B. Hydrolysis of cisplatin–a first-principles metadynamics study. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10348–10355.

[159] Fu, C. F.; Tian, S. X. Molecular dynamics study of solvation differences between cis- and transplatin molecules in water. *J. Chem. Phys.* **2010**, *132*, 174507–174513.

[160] Springer, A.; Bürgel, C.; Böhrsch, V.; Mitric, R.; Bonacic-Koutecky, V.; Linscheid, M. W. The gas-phase chemistry of cis-diammineplatinum(II) complexes: a joint experimental and theoretical study. *ChemPhysChem* **2006**, *7*, 1779–1785.

[161] Dans, P. D.; Coitiño, E. L. Density functional theory characterization and descriptive analysis of cisplatin and related compounds. *J. Chem. Inf. Model.* **2009**, *49*, 1407–1419.

[162] Zimmermann, T.; Chval, Z.; Burda, J. V. Cisplatin interaction with cysteine and methionine in aqueous solution: computational DFT/PCM study. *J. Phys. Chem. B* **2009**, *113*, 3139–3150.

[163] Zhang, Y.; Guo, Z.; You, X. Z. Hydrolysis theory for cisplatin and its analogues based on density functional studies. *J. Am. Chem. Soc.* **2001**, *123*, 9378–9387.

[164] Zimmermann, T.; Zeizinger, M.; Burda, J. V. Cisplatin interaction with cysteine and methionine, a theoretical DFT study. *J. Inorg. Biochem.* **2005**, *99*, 2184–2196.

[165] Carloni, P.; Sprik, M.; Andreoni, W. Key Steps of the cis-Platin-DNA Interaction: Density Functional Theory-Based Molecular Dynamics Simulations. *J. Phys. Chem. B* **2000**, *104*, 823–835.

[166] Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

[167] Car, R.; Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

[168] Laio, A.; VandeVondele, J.; Rothlisberger, U. D-RESP: Dynamically Generated Electrostatic Potential Derived Charges from Quantum Mechanics/-Molecular Mechanics Simulations. *J. Phys. Chem. B* **2002**, *106*, 7300–7307.

[169] Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.* **2002**, *116*, 6941–6947.

[170] Rothlisberger, U,; Carloni P. Simulations of enzymatic systems: Perspectives from Car-Parrinello Molecular dynamics simulations; in Erikson L. A. Theoretical and computational chemistry, Politzer, P.; Maksic, Z. B., Eds.; Elservier: Amsterdam **2001**, Vol. *9*, 215–251.

[171] Dal Peraro, M.; Ruggerone, P.; Raugei, S.; Gervasio, F. L.; Carloni, P. Investigating biological systems using first principles Car-Parrinello molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2007**, *17*, 149–156.

[172] Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed. Engl.* **2009**, *48*, 1198–1229.

[173] Wang J.; Cieplak P.; Kollman P. A. Geometrical and electronic structure variability of the sugar-phosphate backbone in nucleic acids. *J. Comput. Chem.* **2000**, *21*, 1049–1074.

[174] Maurer, P.; Laio, A.; Hugosson, H. W.; Colombo, M. C.; Rothlisberger, U. Automated Parametrization of Biomolecular Force Fields from Quantum Mechanics/Molecular Mechanics (QM/MM) Simulations through Force Matching. *J. Chem. Theory Comput.* **2007**, *3*, 628–639.

[175] Spiegel, K.; Magistrato, A.; Maurer, P.; Ruggerone, P.; Rothlisberger, U.; Carloni, P.; Reedijk, J.; Klein, M. L. Parameterization of azole-bridged dinuclear platinum anticancer drugs via a QM/MM force matching procedure. *J. Comput. Chem.* **2008**, *29*, 38–49.

[176] Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13749–13754.

[177] Chen, K. G.; Valencia, J. C.; Lai, B.; Zhang, G.; Paterson, J. K.; Rouzaud, F.; Berens, W.; Wincovitch, S. M.; Garfield, S. H.; Leapman, R. D.; Hearing, V. J.; Gottesman, M. M. Melanosomal sequestration of cytotoxic drugs contributes to the intractability of malignant melanomas. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9903–9907.

[178] Ishida, S.; Lee, J.; Thiele, D. J.; Herskowitz, I. Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14298–14302.

[179] Aller, S. G.; Unger, V. M. Projection structure of the human copper transporter CTR1 at 6-A resolution reveals a compact trimer with a novel channel-like architecture. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3627–3632.

[180] Larson, C. A.; Adams, P. L.; Jandial, D. D.; Blair, B. G.; Safaei, R.; Howell, S. B. The role of the N-terminus of mammalian copper transporter 1 in the cellular accumulation of cisplatin. *Biochem. Pharmacol.* **2010**, *80*, 448–454.

[181] Wang, X.; Du, X.; Li, H.; Chan, D. S.; Sun, H. The effect of the extracellular domain of human copper transporter (hCTR1) on cisplatin activation. *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 2706–2711.

[182] Puig, S.; Lee, J.; Lau, M.; Thiele, D. J. Biochemical and genetic analyses of yeast and human high affinity copper transporters suggest a conserved mechanism for copper uptake. *J. Biol. Chem.* **2002**, *277*, 26021–26030.

[183] Wang, X.; Guo, Z. The role of sulfur in platinum anticancer chemotherapy. *Anticancer Agents Med. Chem.* **2007**, *7*, 19–34.

[184] Brabec, V. DNA modifications by antitumor platinum and ruthenium compounds: their recognition and repair. *Prog. Nucleic Acid Res. Mol. Biol.* **2002**, *71*, 1–68.

[185] Zhang, W.; Hou, T.; Schafmeister, C.; Ross, W. S.; Case, D. A. LEaP and gleap, AmberTools User's Manual, Ver. 1.5, **2011**.

[186] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

[187] Micheletti, C.; Seno, F.; Maritan, A. Recurrent oligomers in proteins: an optimal scheme reconciling accurate and concise backbone representations in automated folding and design studies. *Proteins: Structure, Function, and Bioinformatics* **2000**, *40*, 662–674.

[188] CPMD, version 3.15.1; IBM Corp. 1990-2008, MPI für Festkörperforschung Stuttgart 1997-2001, 2011; http://www. cpmd.org.

[189] van Gunsteren, W. F. Biomolecular Simulation: The GROMOS96 Manual and User Guide; Hochschulverlag AG der ETH Zurich: Zurich, **1996**.

[190] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.

[191] Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958.

[192] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

[193] Smith, D. E.; Dang, L. X. Computer simulations of NaCl association in polarizable water. *J. Chem. Phys.* **1994**, *100*, 3757–3766.

[194] Joung, I. S.; Cheatham,III, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

[195] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[196] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

[197] Troullier, N.; Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **1991**, *43*, 1993–2006.

[198] Martyna, G. J.; Tuckerman, M. E. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.* **1999**, *110*, 2810–2821.

[199] Nose, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511–519.

[200] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* **1985**, *31*, 1695–1697.

[201] Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2633.

[202] Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45*, 1196–1199.

[203] Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

[204] Kabsch, W.; Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637.

[205] Joosten, R. P.; Te Beek T. A. H.; Krieger, E.; Hekkelman, M. L.; Hooft, R. W. W.; Schneider, R.; Sander, C.; Vriend, G. A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **2011**, *39*, D411–D419.

[206] Bulheller, B. M.; Hirst, J. D. DichroCalc–circular and linear dichroism online. *Bioinformatics* **2009**, *25*, 539–540.

[207] Besley, N. A.; Hirst, J. D. Theoretical Studies toward Quantitative Protein Circular Dichroism Calculations. *J. Am. Chem. Soc.* **1999**, *121*, 9636–9644.

[208] Putrino, A. ; Sebastiani, D. ; Parrinello, M. Generalized variational density functional perturbation theory. *J. Chem. Phys.* **2000**, *113*, 7102–7109.

[209] Sebastiani, D.; Parrinello, M. A New ab-Initio Approach for NMR Chemical Shifts in Periodic Systems. *J. Phys. Chem. A* **2001**, *105*, 1951–1958.

[210] Röhrig. U. F.; Sebastiani, D. NMR Chemical Shifts of the Rhodopsin Chromophore in the Dark State and in Bathorhodopsin: A Hybrid QM/MM Molecular Dynamics Study. *J. Phys. Chem. B* **2008**, *112*, 1267–1274.

[211] te Velde. G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

[212] Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an order-N DFT method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.

[213] ADF2010, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com.

[214] Hammer, B.; Hansen, L. B.; Norskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **1999**, *59*, 7413–7421.

[215] Sterzel, M.; Autschbach, J. Toward an accurate determination of 195Pt chemical shifts by density functional computations: the importance of unspecific solvent effects and the dependence of Pt magnetic shielding constants on structural parameters. *Inorg. Chem.* **2006**, *45*, 3316–3324.

[216] van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597–4610.

[217] Wolff, S. K.; Ziegler, T.; van Lenthe, E.; Baerends, E. J. Density functional calculations of nuclear magnetic shieldings using the zeroth-order regular approximation (ZORA) for relativistic effects: ZORA nuclear magnetic resonance. *J. Chem. Phys.* **1999**, *110*, 7689–7698.

[218] Rehr, J. J.; Mustre de Leon, J.; Zabinsky, S. I.; Albers, R. C. Theoretical x-ray absorption fine structure standards. *J. Am. Chem. Soc.* **1991**, *113*, 5135–5140.

[219] Rehr, J. J.; Albers, R. C.; Zabinsky, S. I. High-order multiple-scattering calculations of x-ray-absorption fine structure. *Phys. Rev. Lett.* **1992**, *69*, 3397–3400.

[220] Rehr, J. J. ; Albers, R. C. Theoretical approaches to x-ray absorption fine structure. *Rev. Mod. Phys.* **2000**, *72*, 621–654.

[221] Newville, M.; Ravel, B.; Haskel, D.; Rehr, J. J.; Stern, E. A.; Yacoby, Y. Analysis of multiple-scattering XAFS data using theoretical standards. *Physica B* **1995**, *208*, 154–156.

[222] Korbas, M.; Marsa, D. F.; Meyer-Klaucke, W. KEMP: A program script for automated biological x-ray absorption spectroscopy data reduction. *Rev. Sci. Instrum.* **2006**, *77*, 063105–063109.

[223] Perczel, A.; Hollosi, M.; Foxman, B. M.; Fasman, G. D. Conformational analysis of pseudocyclic hexapeptides based on quantitative circular dichroism (CD), NOE, and x-ray data. The pure CD spectra of type I and type II .beta.-turns. *J. Am. Chem. Soc.* **1991**, *113*, 9772–9784.

[224] Hess, B. Convergence of sampling in protein simulations. *Phys. Rev. E* **2002**, *65*, 031910–031919.

[225] Zabinsky, S. I.; Rehr, J. J.; Ankudinov, A.; Albers, R. C.; Eller, M. J. Multiple-scattering calculations of x-ray-absorption spectra. *Phys. Rev. B* **1995**, *52*, 2995–3009.

[226] Best, R. B.; Buchete, N.-V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys. J.* **2008**, *95*, L07–L09.

[227] Thompson, E. J.; DePaul, A. J.; Patel, S. S.; Sorin, E. J. Evaluating molecular mechanical potentials for helical peptides and proteins. *PLoS One.* **2010**, *5*, e10056.

[228] Best, R. B.; Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

[229] Best, R. B.; Mittal, J. Balance between alpha and beta structures in ab initio protein folding. *J. Phys. Chem. B* **2010**, *114*, 8790–8798.

[230] Machonkin, T. E.; Westler, W. M.; Markley, J. L. Paramagnetic NMR spectroscopy and density functional calculations in the analysis of the geometric and electronic structures of iron-sulfur proteins. *Inorg. Chem.* **2005**, *44*, 779–797.

[231] Liptak, M. D.; Wen, X.; Bren, K. L. NMR and DFT investigation of heme ruffling: functional implications for cytochrome c. *J. Am. Chem. Soc.* **2010**, *132*, 9753–9763.

[232] Chan, J.; Merrifield, M. E.; Soldatov, A. V.; Stillman, M. J. XAFS spectral analysis of the cadmium coordination geometry in cadmium thiolate clusters in metallothionein. *Inorg. Chem.* **2005**, *44*, 4923–4933.

[233] Solomon, E. I. Spectroscopic methods in bioinorganic chemistry: blue to green to red copper sites. *Inorg. Chem.* **2006**, *45*, 8012–8025.

[234] Csaszar, P.; Pulay, P. Geometry Optimization by DIIS. *J. Mol. Struct., THEOCHEM* **1984**, *114*, 31–34.

[235] Ryckaert, J. P.; Ciccotti, G.; Berendsen H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys. 1977*, *23*, 327–341.

[236] Cheatham, T. E. III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Protein. *J. Am. Chem. Soc.* **1995**, *117*, 4193–4194.

[237] Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.

[238] Jameson, A. K.; Jameson, C. J. Gas phase 13C chemical shifts in the zero-pressure limit: Refinements to the absolute shielding scale for 13C. *Chem. Phys, Lett.* **1987**, *134*, 461–466.

[239] Gregor, T.; Mauri, F.; Car, R. A comparison of methods for the calculation of NMR chemical shifts. *J. Chem. Phys.* **1999**, *111*, 1815–1822.

[240] Macke T. J., *et al.*, AmberTools User's Manual, Ver. 1.3, **2010**.

[241] Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Raghavachari, K.; Schlegel, H. B.; Robb, M.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J. R.; Whiteside, A.; Seeger, Melius, C. F.; Baker, J.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; and Pople, J. A. Gaussian 90 (Gaussian, Inc., Pittsburgh, PA, **1990**).

[242] Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

[243] Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[244] Parr, R. G.; Yang, W. Density-functional theory of atoms and molecules. Oxford University Press, US, **1989**.

[245] Hohenberg, P.; Kohm, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.

[246] Koch, W.; Holthausen, M. C. A chemist's guide to Density Functional Theory. 2nd Ed., Wiley-VCH: Germany, **2001**

[247] Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

[248] Perdew, J. P.; Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **1992**, *45*, 13244–13249.

[249] Heine, V. The concept of Pseudopotentials. *Solid state physics* **1970**, *24*, 1–36.

[250] Ihm, J. Total energy calculations in solid state physics. *Rep. Prog. Phys.* **1988**, *51*, 105–142.

[251] Pickett, W. E. Pseudopotential methods in condensed matter applications. *Comp. Phys. Rep.* **1989**, *9*, 115–197.

[252] Phillips, J. C.; Kleinman, L. New Method for Calculating Wave Functions in Crystals and Molecules. *Phys. Rev.* **1959**, *116*, 287–294.

[253] http://www.tcm.phy.cam.ac.uk/ pdh1001/thesis/node19.html#sec:pseudo

[254] Mackerell, A. D. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *et al. J. Chem. B* **1998**, *102*, 3586–3616.

[255] Allen, M. P.; Tildesley, D. J. Computer simulation of liquids. Oxford University Press, USA, **1989**.

[256] Frenkel, D.; Smit, B. Understanding Molecular Simulation: From Algorithms to Applications. Academic Press, USA, **2001**.

[257] Marx, D.; Jurg, H. Ab initio molecular dynamics: Basic Theory and advanced methods. Cambridge University Press, UK, **2009**.

[258] Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061–1071.

[259] Tully, J. C. Mixed quantum-classical dynamics: Mean-field and surface-hoppin. In Berne, B. J.; Ciccotti, G.; Coker, D. F. Eds. Classical and quantum dynamics in condensed phase simulations. World Scientific, Singapore, **1998**.

[260] Dirac, P. A. M. The principles of Quantum mechanics. Oxford University Press, UK, **1947**.

[261] Sambe, H.; Felton, R. H. A new computational approach to Slater's SCF-Xα equation. *J. Chem. Phys.* **1975**, *62*, 1122–1126.

[262] Goldstein, H.; Poole, C. P.; Safko, J. L. Classical mechanics. Addison-Wesley, USA, **2002**.

[263] Ewald, P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253–287.

[264] van Gunsteren, W. F.; Berendsen, H. J. C. Computer Simulation of Molecular Dynamics: Methodology, Applications and Perspectives in Chemistry. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992–1023.

[265] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

[266] Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393.

[267] Swenden, R. H.; Wang, J.-S. Replica Monte Carlo Simulation of Spin-Glasses. *Phys. Rev. Lett.* **1986**, *57*, 2607–2609.

[268] Hukushima, K.; Nemoto, K. Exchange Monte Carlo Method and Application to Spin Glass Simulations. *J. Phys. Soc. Jpn.* **1996**, **65**, 1604–1608.

[269] Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

[270] Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.* **2002**, *116*, 9058–9067.

[271] Patriksson, A.; van der Spoel, D. A temperature predictor for parallel tempering simulations. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2073–2077.

[272] Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta.* **1977**, *44*, 129–138.

[273] Helgaker, T.; Jaszunski, M.; Ruud, K. Ab Initio Methods for the Calculation of NMR Shielding and Indirect Spin-Spin Coupling Constants. *Chem. Rev.* **1999**, *99*, 293–352.

[274] Keith, T. A.; Bader, R. F. W. *Chem. Phys. Lett.* **1993**, *210*, 223–231.

[275] Bunker, G. Introduction to XAFS: A practical guide to X-ray Absorption fine structure Spectroscopy. Cambridge University Press, UK (**2010**).

[276] Rehr, J. J.; Booth, C. H.; Bridges, F.; Zabinsky, S. I. X-ray-absorption fine structure in embedded atoms. *Phys. Rev. B* **1994**, *49*, 12347–12350.

[277] Beni, G.; Platzman, P. M. Temperature and polarization dependence of extended x-ray absorption fine-structure spectra. *Phys. Rev. B* **1976**, *14*, 1514–1518.

[278] Rosenfeld, L. *Z. Phys.* **1928**, *52*, 161–174.

[279] Bayley, P. M.; Nielsen, E. B.; Schellman, J. A. Rotatory properties of molecules containing two peptide groups: theory. *J. Phys. Chem.* **1969**, *73*, 228–243.

[280] Bulheller, B. M.; Rodger, A.; Hirst, J. D. Circular and linear dichroism of proteins. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2020–2035.

[281] Wallace, B. A.; Janes, R. W. Modern Techniques for Circular Dichroism and Synchrotron Radiation Circular Dichroism Spectroscopy. IOS Press, Netherlands, (**2009**).

[282] Woody, R. W.; Sreerama, N. Comment on "Improving protein circular dichroism calculations in the far-ultraviolet through reparametrizing the amide chromophore" [J. Chem. Phys. 109, 782 (1998)]. *J. Chem. Phys.* **1999**, *111*, 2844–2845.

[283] Atkins, P.; de Paula, J. Physical Chemistry. 9th Ed., W. H. Freeman, **2009**.

[284] Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566.

[285] Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* **2008**, *71*, 126601–126622.

[286] Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603–020606.

[287] Bussi, G.; Gervasio, F. L.; Laio, A.; Parrinello, M. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J. Am. Chem. Soc.* **2006**, *128*, 13435–13441.

[288] Bonomi, M.; Parrinello, M. Enhanced Sampling in the Well-Tempered Ensemble. *Phys. Rev. Lett.* **2010**, *104*, 190601–190604.

[289] Cheatham, T. E. III Simulation and modeling of nucleic acid structure, dynamics and interactions. *Curr. Opin. Struct. Biol.* **2004**, *14*, 360–367.

[290] Rohs, R.; West, S. M.; Liu, P.; Honig, B. Nuance in the double-helix and its role in protein-DNA recognition. *Curr. Opin. Struct. Biol.* **2009**, *19*, 171–177.

[291] Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E. III; Laughton, C. A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys J.* **2007**, *92*, 3817–3829.

[292] Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* **2009** *37*, 5917–5929.

[293] http://www.prace-project.eu/

List of Figures

1.1 Chemical structures of a) cisplatin (*cis*-diamminedichloroplatinum(II)) and b) transplatin (*trans*-diamminedichloridoplatinum(II)). 1

1.2 A simplified scheme of the mechanisms affecting cellular uptake, distribution and efflux of cisplatin (adapted from [30]). 3

2.1 Major cisplatin-DNA adducts: (a) 1,2-$[Pt(NH_3)_2]^{2+}$-d(GpG) intra-strand cross-link; (b) 1,3-$[Pt(NH_3)_2]^{2+}$-d(GpTpG) intra-strand cross-link; (c) inter-strand cross-link (adapted from [24]). DNA's backbone is shown in ribbon; their nucleobases in bond representation; the platinum atom is represented as a sphere. 7

2.2 Structures of cisplatin-DNA adduct: (a) from X-ray [41] and (b) from NMR [44]. Atoms in DNA are represented by lines while the cisplatin moiety is shown in ball-stick model (platinum in gray, nitrogen in blue). 8

2.3 X-ray structure of the complex between *HMGB1* domain A and *cis*-[Pt(II)$(NH_3)_2$]-d-(CCUCTCTG*G*ACCTTCC)-d(GGAGAGACCTGGAAGG) (PDB code 1CKT) [76]. The DNA's backbone is shown in blue ribbon; nucleobases in blue bonds; cisplatin moiety in van der Waals balls. The *HMGB1*'s backbone is shown in red ribbon; its Phe37 amino acid, which inserts between two platinated G8 and G9 bases, is shown in green ball-stick representation. 9

2.4 Correlation between the fold-resistance (which is defined as the fold increase in the number of survival cells when treated with the same amount of drug) and the percent decrease in cisplatin accumulation with respect to sensitive cells (adapted from [30]). 10

2.5 Cartoon representing the copper influx transporter *Ctr1* (adapted from [145]). Left: *Ctr1* is a symmetrical homotrimeric trans-membrane protein. Nine trans-membrane helices form a pore in the lipid bilayer. Each subunit is represented by a different color. Right: One of the three subunits is removed to reveal the interior of the pore. 13

2.6 Overlay of 2D ^{1}H,^{13}C-edited HSQC spectra of *Mets7* (red contours) and of mixtures (1:1) of *Mest7* with a) cisplatin and b) transplatin. Spectral changes induced by cisplatin/transplatin are indicated with arrows. Figures adapted from [142]. 15

2.7 The NMR structures of: (a) *Atox1* [151] and (b): the soluble domain 4 of the *ATP7B* protein [148]. The backbone is represented with ribbon. Different colors indicate different secondary-structure types: green is loop, yellow is beta sheet, and purple is alpha helix. Two cysteine amino acids are shown in bond representation. 16

3.1 Schematic diagram of the relationship between all-electron and pseudopotentials and wavefunctions as well as real and pseudopotentials, adapted from [253]. 29

3.2 Schematic representation of Coulomb and modified Coulomb potentials. 31

4.1 Structural models of platinated and Apo peptides emerging from REST and QM/MM simulations. The backbone conformations of the last QM/MM snapshot (in red ribbons) and a representative REST snapshot (in blue ribbons) are superimposed. The QM atoms are shown in stick-and-balls (CB=C_β, CG=C_γ, CE=C_ϵ). Water is shown in lines. Counterions are not shown for the sake of clarity. 52

4.2 Potential energy distribution of 8 replicas (in different colors) simulated with REST method for the platinated peptides C1-C3 and T1-T3 and the apopeptide. 55

4.3 Running averages of CD spectra calculated from the REST trajectories for cisplatn-*Mets7* (C1-C3) and transplatin-*Mets7* (T1-T3). . . . 56

4.4 Calculated and experimentally measured [142] CD spectra of cisplatin-*Mets7* (a) and transplatin-*Mets7* (b) adducts. 60

4.5 Pt-O radial pair distribution functions calculated from the QM/MM trajectories. Left: cisplatin-*Mets7*, right: transplatin-*Mets7*. 60

4.6 Calculated and experimental EXAFS spectra at the Pt edge of cisplatin-*Mets7* (a) and transplatin-*Mets7* (b) adducts. The experimental spectrum has been measured in this work. Fourier Transforms of the spectra in (a) and (b) are given in (c) and (d), respectively. They are plotted as a function of the scattering path length of the photoelectron R. Note that R does not correspond to a real interatomic distance since the peak positions are shifted inward, with respect to the bond distances; due to the phase shifts of the photoelectron (i.e. increase in its kinetic energy caused by the electronic potentials of both the scattering and the absorbing atoms) [225]. The shift, therefore, depends on both the absorbing and scattering atom types, and in our cases is approximately 0.4 Å (Tab. 4.5. 62

4.7 Contributions from different types of scattering atoms to the total EXAFS spectra, calculated from the QM/MM trajectories. 67

5.1 The X-ray structure of cisplatin-DNA-*HMGB1*. The DNA is shown with bond representation, HMGB1 with Connolly surface representation and cisplatin moiety in van der Waals balls. Hydrogen atoms missing in the X-ray structure were added using the Leap program [185]. 70

5.2 Complex of platinated DNA-*HMGB1*. The backbone conformations of X-ray structure (red), the QM/MM structure (blue) and the classical MD structure (green) are superimposed. Platinated guanines are shown in ball-and-stick. 73

5.3 The backbone RMSD of QM/MM trajectory. 74

5.4 Atom name convention used for parameterization of partial charges. . 77

5.5 Atom type convention used for parameterization of bonded parameters. 77

5.6 The backbone RMSD of classical MD trajectory. 78

5.7 Number of hydrogen bonds between the platinated DNA and *HMGB1* 78

5.8 The contact map, which is defined as matrix of the smallest distances between pairs of residues, of the X-ray structure (a) and the classical MD trajectory (b). 79

A.1 DFT optimized model compounds 87

A.2 Atom name and atom type convention for the QM atoms. The atom names are in bold and atom types in italics. Hydrogen atoms are not shown for the sake of clarity 94

A.3 RMSD of the QM region during the Car-Parrinello QM/MM simulations. Left: cisplatin-*Mets7*, right: transplatin-*Mets7*. 94

List of Tables

4.1 Selected features of the simulated models 53

4.2 The averaged RMSD of the REST structures with respect to the corresponding last QM/MM structures. 58

4.3 The root mean square fluctuations (RMSF), which is defined as the standard deviation of the backbone RMSD. The smaller the value of RMSF, the more rigid the conformation, and vice versa. 59

4.4 DSSP secondary structure assignment for the ensembles of conformations generated by REST simulations. 59

4.5 Structural parameters calculated from QM/MM and REST simulations for cisplatin-*Mets7* (C1-C3) and transplatin-*Mets7* (T1-T3) adducts. Standard deviations are shown in parentheses. Bond lengths are in angstrom and angles in degree. 61

4.6 ^{13}C (top) and ^{1}H (bottom) NMR chemical shifts (in ppm) of the Pt-coordinated Met residues of cisplatin-*Mets7* (C1–C3) and transplatin-*Mets7* (T1–T3) adducts after REST simulations. The corresponding values before REST are reported in Tables. A.8–A.10 in Appendix. Standard deviations are shown in parentheses. See Fig. 4.1 for atom name conventions. 64

4.7 ^{195}Pt NMR chemical shifts (in ppm) of the cisplatin-*Mets7* (C1–C3) and transplatin-*Mets7* (T1–T3) adducts after REST simulations. The corresponding values before REST are reported in Tab. A.11 in Appendix. Standard deviations are shown in parentheses. 64

5.1 Selected features of the qmmm simulation box. 72

5.2 Comparison of the selected structural features of QM/MM, classical MD models with those of the X-ray structure. See Fig. 5.5 for the atom type convention. 72

5.3 Partial charges (in fraction of electronic charge) derived from the force matching procedure for the cisplatin-DNA-*HMGB1* complex. See Fig. 5.4 for the atom name convention. 75

5.4 The harmonic force constant K_r (kcal/mol Å^{-2}), K_θ (kcal/mol rad^{-2}) of bond stretching and angle bending, respectively, derived from the force matching procedure for the cisplatin-DNA-*HMGB1* complex. r_{eq} (Å) and θ_{eq} (degree) are equilibrium bond lengths and bond angles, respectively. See Fig. 5.5 for the atom type convention. 76

5.5 Force constant of dihedral angles (kcal/mol) obtained from the force matching procedure. See Fig. 5.5 for the atom type convention. . . . 76

A.1 Structural parameters of the model compounds optimized at the DFT level in gas phase. S_{cis} and S_{trans} are sulphur atoms cis and trans to the fourth ligand (chloride, water molecule, or OH^- group). 88

A.2 Charges (in fraction of electronic charge) derived from the force matching procedure for the cisplatin-*Mets7* adduct models. See Fig. A.2 for the atom name convention. 88

A.3 Charges (in fraction of electronic charge) derived from the force matching procedure for the transplatin-*Mets7* adduct models. See Fig. A.2 for the atom name convention. 89

A.4 The harmonic force constant K_r of bond stretching (kcal/mol $Å^{-2}$) and equilibrium bond lengths r_{eq} (Å) obtained from the force matching procedure. See Fig. A.2 for the atom type convention. 89

A.5 The harmonic force constant of angle bending (kcal/mol rad^{-2}) and equilibrium bond angles (degree) obtained from the force matching procedure. See Fig. A.2 for the atom type convention. 90

A.6 Force constant of dihedral angles (kcal/mol) obtained from the force matching procedure. See Fig. A.2 for the atom type convention. . . . 91

A.7 The cosine content of the first four principal components of the QM region in the QM/MM simulations. 91

A.8 NMR chemical shifts of C1-C3 (^{13}C (top) and ^{1}H (bottom)) obtained from QM/MM simulations before REST simulations. 92

A.9 NMR chemical shifts of T1-T3 (^{13}C (top) and ^{1}H (bottom)) obtained from QM/MM simulations before REST simulations. 92

A.10 NMR chemical shifts of Apo (^{13}C (top) and ^{1}H (bottom)) obtained from QM/MM simulations for the representative snapshot of the REST trajectory. The RMSD of this representative snapshot relative to the last REST one is 3.3 Å. 93

A.11 ^{195}Pt NMR chemical shifts obtained from QM/MM simulations before REST simulations. 93

Curriculum Vitae

Personal Data

Last name:	Nguyen
First name:	Trung Hai
Date of birth:	October 2nd, 1983
Place of birth:	An Giang, Vietnam
Nationality:	Vietnamese
Email:	n.t.hai@grs-sim.de; nguyentrungh@gmail.com

Qualifications

2001	High school diploma at Tran Van Thanh High School, An Giang, Vietnam.
2001–2005	B.Sc. in physics at Vietnam National University, Ho Chi Minh City, Vietnam. *Thesis*: "Studying the vortex glass transition in disordered type-II superconductors using the method of finite-size scaling analysis of I-V characteristics." *Supervisor*: Assoc. Prof. Dr. Hoang Zung.
2007–2009	M.Sc. at the Graduate School of Science, Osaka University, Japan. *Study field*: Condensed Matter Physics. *Thesis*: "Ordering of the pyrochlore Heisenberg antiferromagnet with the ferromagnetic next-nearest-neighbor interaction." *Supervisor*: Prof. Dr. Hikaru Kawamura.
Since Dec. 2009	Doctoral candidate in the Computational Biophysics group of German Research School for Simulation Sciences GmbH, Jülich, Germany. *Thesis*: "Action of the anticancer drug cisplatin investigated by molecular simulation." *Supervisor*: Univ.-Prof. Ph.D. Paolo Carloni.