From All-Atom Molecular Mechanics to Coarse-Grained Lattice Models: Computational Approaches to Problems in Protein Biochemistry

by

John Peter Cvitkovic

A Dissertation

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Doctor of Philosophy

in

Chemistry

May 2019

Abstract

Computational simulations of chemical systems play an ever-increasing role in many areas of biochemical research from rational drug design to probing fundamental physiological processes. Depending on the method, a vast array of properties are able to be predicted. Here we report the design and implementation of two methods for investigating diverse problems in protein biochemistry.

In order to better understand protein–metal interactions—most importantly for the difficult to model transition metal ions—empirical force field parameters were developed for Pt(II), cisplatin, and other Pt(II) coordination compounds. Two force field frameworks were used: a modified version of the fixed-charge OPLS-AA and the polarizable POSSIM force field. A seven-site model was used for the Pt(II) ion. The produced parameters are compatible with the OPLS-AA and POSSIM force fields and can be used in protein–metal binding simulations in which—contrary to the common treatment of metal ions in such simulations—the position or even coordination of the ion does not have to be constrained using preexisting knowledge. It has been demonstrated that the produced models are capable of reproducing key properties of relevant Pt(II) complexes but that the POSSIM formalism yields more accurate values for energies of formation than the OPLS-AA model.

This Pt(II) model was employed—along with previously developed Cu(I) parameters—to investigate the binding of platinum to the protein Atox1, a human copper chaperone implicated in the resistance mechanism of cisplatin and other platinum antitumor compounds. In collaboration with the Dmitriev and Bernholc groups, we used our models to inform and refine spectroscopic experiments as well as to serve as starting points for high-performance quantum calculations. It was shown that under physiological redox conditions, copper(I) and cisplatin can form large polymers with glutathione. These polymers were capable of transferring copper(I) to *apo*-Atox1 or to platinum(II) to copper-loaded Atox1. Analysis of the simultaneous binding of copper(I) and platinum(II) to Atox1 was found to occur through the formation of copper–sulfur–platinum bridges, where copper is coordinated by three sulfur atoms and platinum by four sulfur atoms.

With the goal of using a simple model to be able to quickly estimate the acid disassociation constants of proteins, PKA17 has been developed and tested. PKA17 is a coarse-grain grid-based method and software tool for accurately and rapidly calculating protein pK_a values given an input PDB structure file. During development, parameter fitting was carried out using a compilation of 442 Asp, Glu, His, and Lys residues that had both high-resolution PDB structures and published experimental pK_a values available. Applying our PKA17 model, the calculated average unsigned error and RMSD for the residue set were found to be 0.628 and 0.831 pH units, respectively. As a benchmark for comparison, the same residue set was evaluated with the PROPKA software package which resulted in an average unsigned error of 0.761 pH units and an RMSD of 1.063 pH units. Finally, a web interface for the PKA17 software was developed and deployed (http://users.wpi.edu/~jpcvitkovic/pka_calc.html) to make PKA17 available to the wider scientific community.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction and Background

Computer simulations of chemical systems play an increasing role in many areas of biochemical research from rational drug design to the understanding of fundamental biochemical processes. Depending on the method, computer simulations are able to predict a vast array of properties from reaction kinetics to free energies of binding. In particular, computer simulations are key in exploring properties not easily accessed by experiment for example, the dynamics and microstructure of solvated systems. The key factor in choosing a general method is the size of the system of interest. For example, quantum mechanical methods are able to very accurately reproduce many disparate properties but for systems containing more than a few hundred atoms, the calculation become prohibitively expensive even on large parallelized computing clusters. For large systems containing up to many thousands of atoms, such as those containing proteins or other biological macromolecules, empirical force fields offer a computationally affordable alternative.

However, the use of chemical force fields is not without drawbacks. Many force fields neglect explicit electrostatic polarization because such calculations can be computationally demanding and because this simplifies the parametrization process. Even without explicit polarization, force fields have been rather successful in modeling a variety of systems including those containing protein-small molecule interactions. However, some systems, such as those containing small molecule–ion (***1***,***2***) and protein–ion (***3***) interactions are not adequately reproduced without including explicit polarization. Taking this into consideration, we present two models, a polarizable and a non-polarizable model, of cisplatin for use in protein simulations.

Cisplatin and Cellular Resistance to Platinum-Based Therapeutics

Cisplatin, or *cis*-dichlordiammineplatinum(II), is the prototypical member of a group of platinum-based anticancer compounds collectively known as platins which have important clinical roles in the treatment of a variety of cancers (***4***). The structure of cisplatin, its hydrolysis products, and the pK_a values for the key deprotonation steps are shown in Figure 1-1.

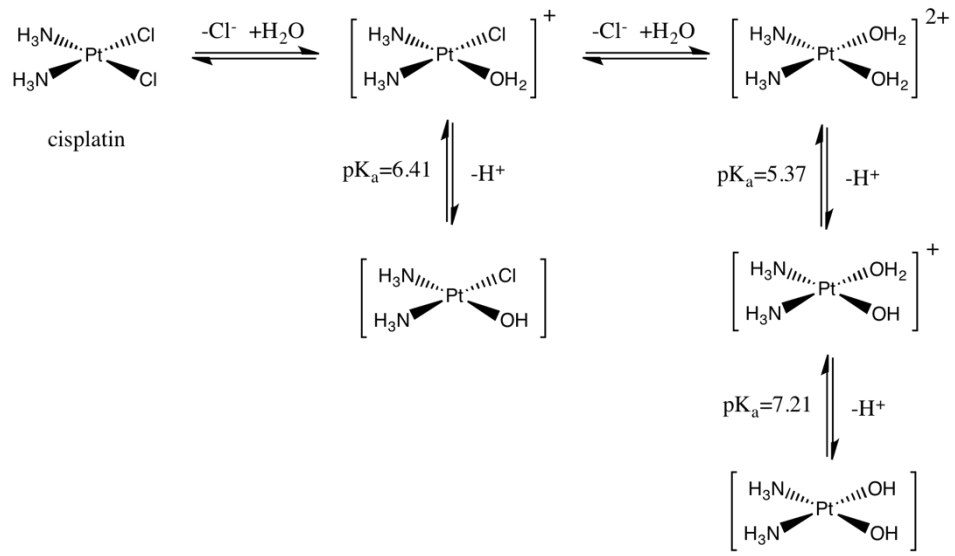


Figure 1-1. Cisplatin, its hydrolysis products, and key associated acid–base equilibria; pK_a values were taken from (***5***).

It is generally agreed that cisplatin's mechanism of anticancer action involves the binding of one of its hydrolysis products to DNA, eventually triggering apoptosis (***4***). Figure 1-2 shows the first hydration product, the thermodynamically favored coordination site (N^7 of guanidine), and the resulting possible DNA adducts. These crosslinked DNA adducts trigger various signaling pathways including various DNA damage recognition mechanisms. While the final cellular outcome is usually apoptosis, the exact pathway from platinum binding to cell death has not been completely elucidated (***4***)

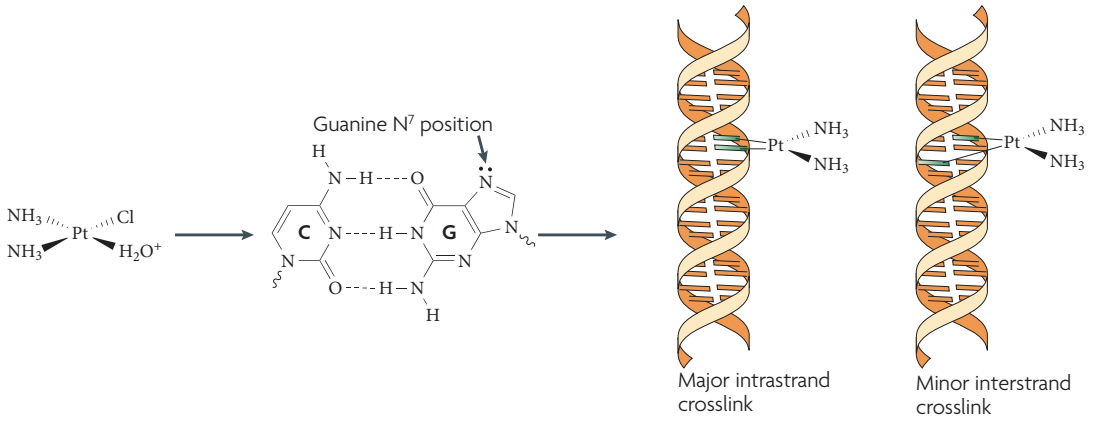


Figure 1-2. General mechanism of cisplatin's anticancer action. First cisplatin is "activated" by the replacement of one or more chloride ligands with water (Figure 1-1). This activated product then coordinates to a DNA fragment with a nitrogen atom on a nucleobase displacing the first water ligand. The N7 of guanidine is most favored but other binding modes are possible. If applicable, the remaining chloride ligand is then replaced by a water molecule from solution. The second water ligand is then replaced by another nitrogen atom on a nucleobase. This forms one of two types of crosslink products: intrastrand and interstrand crosslinks (***4***)**.**

While the initial clinical response to cisplatin is high, most patients will eventually develop reduced response, or resistance, to the drug (***4***). There is a strong body of evidence indicating that the mechanism of cisplatin resistance is due to protein binding and/or protein-mediated transport (***6***) (***7***) Of particular interest is the binding of cisplatin to ATP7A, ATP7B, and/or Atox1 which are two human copper transporters and an associated copper chaperone. Binding to these proteins are of interest because they well-studied (both in their endogenous role and their cisplatin-binding ability) but most importantly because two crystal structures of Atox1 with bound platinum complexes are known (***8***). With this in mind, we are developing two models of cisplatin, a non-polarizable model within the OPLS (***9***) framework and a polarizable model within the POSSIM (***10***) framework, for use in protein simulations.

Fast Predictions of Acid Dissociation Constants in Proteins

Assessing the values of protein acid dissociation constants is important for predicting the structure stability, reactivity, and potential protein–protein interactions. To this end, we have recently developed a simple scheme—PKA17—for accurately and rapidly calculating protein pK_a values given only an input PDB structure file. We have carried out parameter fitting using a total of 442 Asp, Glu, His, and Lys residues for which high quality experimental pK_a values are available. We then compared the performance of PKA17 to several commonly used pK_a prediction methodologies. Three sets of experimental pK_a values were used in this benchmarking: an initial fitting set, a separate test set, and the combined fitting and testing sets. The accuracy of our grid-based methodology compared favorably to the other evaluated pK_a prediction methods. Most importantly, this accuracy did not come at the cost of speed as our grid-based methodology is computationally efficient and almost 10-fold faster than the next fastest surveyed method.

Over the course of our work, we've developed an internet-based version of our software (http://users.wpi.edu/~jpcvitkovic/pka_calc.html) in which any user can upload a PDB file for analysis. Our method is more accurate than the widely used PROPKA calculator and fast enough that—even for the largest proteins—the results are displayed in the user's browser in only a few seconds.

Chapter 2: Force Fields and Water Models

Force Field Background

The OPLS-AA force field and the polarizable POSSIM force field were used for the non-polarizable and polarizable models of cisplatin, respectively. All calculations were carried out with the POSSIM 2.0 software package (***11***). The Optimized Potentials for Liquid Simulations All Atom (OPLS-AA) force field is a fixed-charge force field that explicitly considers every atom and their pairwise interactions. The OPLS force field was developed and validated by Prof. Jorgensen's group at Yale University (***9***). For a given configuration of a molecular ensemble, the total energy, E_{tot}, is calculated as the sum of the following energy components:

$$E_{tot} = E_{stretch} + E_{bend} + E_{torsion} + E_{vdW} + E_{electrostatic} \tag{1}$$

The $E_{stretch}$ and E_{bend} are the harmonic bond stretching and angle bending terms and are given by the following equations:

$$E_{stretch} = \sum_{bonds} K_r\left(r - r_{eq}\right)^2 \tag{2}$$

$$E_{bend} = \sum_{angles} K_\theta\left(\theta - \theta_{eq}\right)^2 \tag{3}$$

where K_r and K_θ are the equilibrium constants for a given bond and angle, r is the bond length, θ is the bond angle, and the *eq* subscripts denote equilibrium values. The summations are performed over each bond and angle in the ensemble. The torsional component of the energy, $E_{torsion}$, is given by the following Fourier series:

$$E_{torsion} = \sum_i \left[\frac{V_{i,1}}{2}(1 + \cos\phi_i) + \frac{V_{i,2}}{2}(1 - \cos 2\phi_i) + \frac{V_{i,3}}{2}(1 + \cos 3\phi_i) + \frac{V_{i,4}}{2}(1 - \cos 4\phi_i)\right] \tag{4}$$

where $V_{i,n}$ are the Fourier coefficients, ϕ_i is a dihedral angle and the summation is performed over all of the dihedral angles, i. The van der Waals energy, E_{vdW}, is given by the standard Lennard-Jones potential (eq 5) using geometric combining rules (eq 6):

$$E_{vdW} = \sum_{i \neq j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^{6} \right] f_{ij} \tag{5}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}; \qquad \sigma_{ij} = \sqrt{\sigma_i \sigma_j} \tag{6}$$

For a pair of atoms i and j, R_{ij} is the distance between the pair, ε_{ij} and σ_{ij} are the Lennard-Jones constants for the pairwise interaction (calculated via eq 6 from the per-atom type constants, ε_i, ε_j, σ_i, and σ_j), and f_{ij} is a scaling factor. The summation is over all i–j atom pairs where $i \neq j$. For pairs of atoms in different molecules, the scaling factor, f_{ij}, is always equal to 1. For pairs of atoms in the same molecule: $f_{ij} = 0$ when the pair is connected by a bond (1–2 pairs) or a bond angle (1–3 pairs), $f_{ij} = 0.5$ when the pair is separated by three bonds (1–4 pairs), and $f_{ij} = 1$ for all other pairs. Finally, the electrostatic part of the total energy, $E_{electrostatic}$, is given by the Coulomb potential:

$$E_{electrostatic} = E_{Coulomb} = \sum_{i \neq j} k_e \frac{q_i q_j}{R_{ij}} f_{ij} \tag{7}$$

where k_e is the Coulomb constant, q_i and q_j are charges of a pair of atoms i and j, and R_{ij} is the distance between the pair. The summation is over all i–j atom pairs where $i \neq j$. The scaling factor, f_{ij}, is determined in the manner as for the van der Waals energy.

The POlarizable Simulations with Second-order Interaction Model (POSSIM) force field is being developed by Prof. Kaminski's group (***11***). POSSIM shares the same functional form as OPLS except for an additional term in the electrostatic component of the total energy to take into account electrostatic polarization:

$$E_{electrostatic} = E_{Coulomb} + E_{pol}; \qquad E_{pol} = -\frac{1}{2} \sum_i \boldsymbol{\mu}_i \mathbf{E}_{i,0} \tag{8}$$

Here, $E_{Coulomb}$ is the Coulomb potential in Equation 7, $\boldsymbol{\mu}_i$ is the induced dipole on the ith polarizable site, and $\mathbf{E}_{i,0}$ is the electric field produced by the permanent charges in the absence of the induced dipoles. An induced dipole, $\boldsymbol{\mu}_i$, on site i, is dependent on that site's isotropic polarizability, α_i, and total electrostatic field $\mathbf{E}_{i,tot}$:

$$\boldsymbol{\mu}_i = \alpha_i \mathbf{E}_{i,tot} \tag{9}$$

$$\mathbf{E}_{i,tot} = \mathbf{E}_{i,0} + \sum_{j \neq i} \mathbf{T}_{ij} \boldsymbol{\mu}_j \; ; \qquad \mathbf{T}_{ij=} \frac{1}{R_{ij}^3} \left(\frac{3 \boldsymbol{R}_{ij} \boldsymbol{R}_{ij}}{R_{ij}^2} - \mathbf{I} \right) \tag{10}$$

where R_{ij} is the distance between atomic sites i and j, and $\mathbf{I}$ the unit tensor. Combining Equations 9 and 10 affords the following expression for the induced dipole:

$$\boldsymbol{\mu}_i = \alpha_i \mathbf{E}_{i,0} + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \boldsymbol{\mu}_j \tag{11}$$

This expression is usually solved iteratively until the change is the induced dipoles are sufficiently small. The first three iterations, giving "first-order," "second-order," and "third-order" approximations for the induced dipoles are as follows:

$$\boldsymbol{\mu}_i^I = \alpha_i \mathbf{E}_{i,0} + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \boldsymbol{\mu}_j^0 = \alpha_i \mathbf{E}_{i,0} \tag{12}$$

$$\boldsymbol{\mu}_i^{II} = \alpha_i \mathbf{E}_{i,0} + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \boldsymbol{\mu}_j^I = \alpha_i \mathbf{E}_{i,0} + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \alpha_j \mathbf{E}_{j,0} \tag{13}$$

$$\begin{aligned} \boldsymbol{\mu}_i^{III} &= \alpha_i \mathbf{E}_{i,0} + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \boldsymbol{\mu}_j^{II} \\ &= \alpha_i \mathbf{E}_{i,0} + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \alpha_j \mathbf{E}_{j,0} + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \alpha_j \sum_{k \neq j} \mathbf{T}_{jk} \alpha_k \mathbf{E}_{k,0} \end{aligned} \tag{14}$$

The first-order approximation has the physical meaning of having inducible dipoles that do not interact with each other, only with the permanent charges. While this approximation is computationally efficient, the improvement over a non-polarizable fixed-charge model is minimal (***11***). The POSSIM force field uses the second-order approximation which does not have a simple physical meaning. The second-order approximation can be thought of as having inducible dipoles with magnitudes determined by their interaction with permanent charges and "first-order" dipoles (*i.e.* dipoles which are only influenced by permanent charges). This second-order approximation is about an order of magnitude faster than explicitly solving Equation 11 as a system of linear equations and results in a negligible loss of accuracy (***11***).

General Fitting Protocols

The procedure for determining values of the potential energy parameters (*i.e.* the constants in Equations 1–9 specific to each atom type or pair of atom types) is an iterative process. First the isotropic polarizabilities are fit using three-body energies (***12***). Next, the permanent charges and Lennard-Jones parameters are fit using gas-phase dimerization energies. Finally, the parameters are fine-tuned by reproducing liquid-phase properties. This is an iterative process because during these steps, small changes in the parameters require reevaluating previous steps to ensure no loss of accuracy. Had stretch, bend, or torsional parameters been needed, they would have been fit to gas-phase energies and geometries and refined using liquid simulations. However, in the systems under study, there are no additional bond stretch, angle bend, or explicitly defined torsions that requiring parametrization.

Atomic Charges and Lennard-Jones Parameters

For small molecules and ions, permanent atomic charges and Lennard-Jones parameters are chosen to best reproduce the quantum mechanical gas-phase dimerization energies and distances. Dimers of the small molecule or ion and water (both with water donating and accepting hydrogen bonds as applicable) as well as homodimers for small molecules are used. Quantum mechanical gas-phase energy minimizations were carried out using the Jaguar electronic structure package (***13***) and the previously developed (***14***) pseudospectral local second-order Møller−Plesset (LMP2) extrapolation method. Briefly, dimer geometries were obtained by LMP2 optimizations using the cc-pVTZ(-f) basis set. Then, fixing the geometry, binding energies were calculated using the cc-pVTZ(-f) basis set and the larger cc-pVQZ(-g) basis set. These energies were then corrected for basis set superposition error (BBSE) using the

counterpoise method. The binding energy of the dimer is then calculated as a linear combination of the BBSE corrected energies (***14***), as shown in Equation 15:

$$E_{bind} = -0.68512\left(E^{BSSE}_{cc-pVTZ(-f)}\right) + 1.68512\left(E^{BSSE}_{cc-pVQZ(-g)}\right) \quad (15)$$

The target dimer distances are taken directly from the LMP2/cc-pVTZ(-f) energy minimizations as the distance between heavy atoms (*e.g.* the N–O distance of $RH_2N\cdots HOH$, an amine–water dimer).

Electrostatic Polarizability

Dipolar probes consisting of two point-charges of opposite sign, magnitude 0.78 *e*, and separation of 0.58 Å were used to apply electrostatic perturbations to a target molecule. The dipole moment of one of these probes is 2.17 D—similar to non-polarizable models of water (***12***). The probes were placed at potential hydrogen bonding sites of the target molecule. The three-body energies were calculated as followed:

$$\begin{aligned} E_{3body} = E(1+2+3) - E(1+2) - E(1+3) - E(2+3) \\ + E(1) + E(2) + E(3) \end{aligned} \quad (16)$$

Figure 2-1 shows the energy components of eq 16 for an example molecule. Note, if there are more than two hydrogen bond sites, there will be a three-body energy for each combination of probes and sites. The isotropic polarizabilities, α_i, were chosen to minimize the root mean squared deviation (RMSD) of the quantum mechanical three-body energies and the POSSIM-calculated three-body energies. The quantum mechanical three-body energies were calculated using density functional theory (DFT), the B3LYP functional , and the cc-pVTZ(-f) basis set (***12***). The calculations were performed using the Jaguar electronic structure package (***13***). In general, this basis set and level of theory gives accurate charge distributions but underestimates the gas-phase polarizability compared to experimental results. Better agreement with experimental gas-phase polarizabilities could be achieved by including diffuse functions to the chosen basis set. However, previous computational studies strongly suggest that fitting to DFT three-body energies including diffuse functions results in an overestimate of the liquid-phase polarizability (***12***). This is likely because in the liquid-phase, Pauli repulsion from nearby molecules increases the energy of the diffuse interactions thereby reducing their contribution to the polarization. Finally, it should be

noted that the three-body energy in a non-polarizable fixed-charge force field is always zero as there are no explicit many-body interactions.

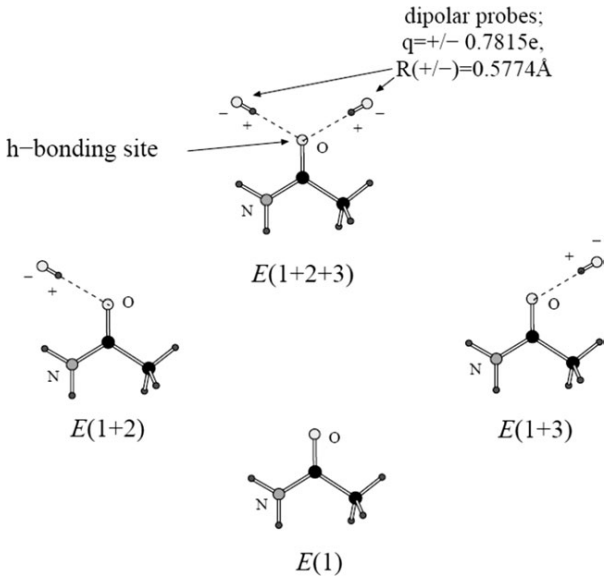


Figure 2-1. Configurations to calculate three-body energies of a small molecule (one three-body energy of acetamide as an example) using dipolar probes. E(2), E(3), and E(2+3) are not shown but are simply the two isolated probes and the two probes in E(1+2+3) without the small molecule (***II***).

Assessing the Second-Order Polarization Model

To illustrate the differences in polarizability approximations, Figure 2-2 shows the polarization energies calculated as a function of distance between two particles—one particle had a charge of +0.5 e, the other a charge of −0.5 *e*, and both with polarizabilities of 2.0 Å^3. The calculations were performed for the full-scale (eq 11), the second-order (POSSIM, eq 13), and the first-order (eq 12) polarizability approaches. In the figure, it can be seen that significant deviations from the full-scale model occur only at small separations.

For example, at a distance of 2.6 Å—a "typical" heavy atom···heavy atom intermolecular bond length—the difference between the full polarization and the second-order model is approximately 5%. Any deviation at shorter (but still physically relevant) distances can be corrected by proper second-order parametrization. In fact, the rapid growth in the magnitude of the full-scale polarization energy is likely to take place in the region where the point dipole approximation is already less valid. For example, at a separation of 1.6 Å, the full polarization energy is −1076 kcal/mol, already too large to be physical.

Note that our second-order POSSIM model does not parametrize the force field for full-scale polarization and then use those parameters in the approximate second-order implementation. All parametrization is done for the second-order polarization approximation, which means that we compensate for any systematic differences between the full-scale and second-order approximations. Even if full-scale polarization were desired, using the second-order POSSIM parameters would likely produce large errors rendering them unusable without significant reparametrization.

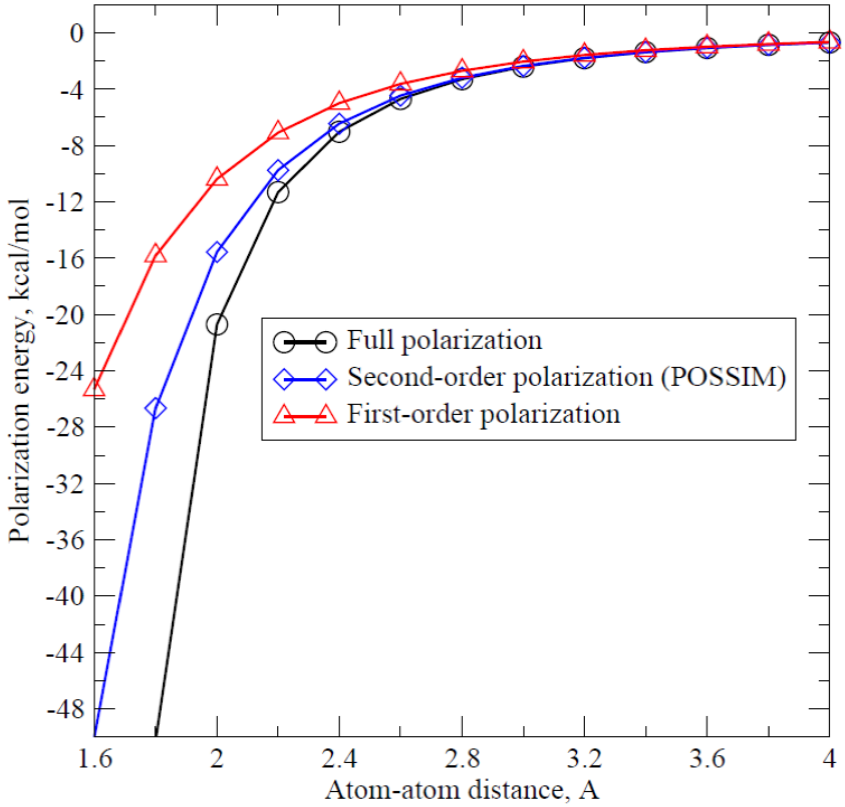


Figure 2-2. Polarization energy between two particles with charges ±0.5 e and polarizability of 2.0 Å^3 as a function of distance between their centers.

In order to further evaluate the second-order polarization model, we have conducted the following experiment (***15***). We equilibrated a simulation of 216 POSSIM water molecules at 1 atm and 25 °C while employing periodic boundary conditions. We then took a single configuration of the system and performed three single point calculations in which only the polarization order was varied.

The first-order polarization energy was found to be −742.12 kcal/mol for this system while the second-order POSSIM formalism yielded a polarization energy of −1109.35 kcal/mol. The full-scale

converged polarization energy was −1729.75 kcal/mol. Thus, the polarization energy calculated per one water molecule was −3.44, −5.14, and −8.01 kcal/mol using the first-order polarization model, POSSIM, and the complete polarization formalism, respectively. Given that our parametrization of the second-order POSSIM model is carried out in such a way as to eliminate possible effects in quantitative differences of magnitudes of polarization energy with the full polarization model, the above numbers do not indicate any fundamental problem with the POSSIM formalism. Rather, the calculations demonstrate the amount of difference with the full-scale polarization methodology that can be encountered when using the second-order approximation.

In addition to examining the energy results for a pure liquid water snapshot, we also ran full-scale Monte Carlo simulations with the three polarization models (***15***). We took a starting configuration from a previously equilibrated periodic box of 216 POSSIM water molecules and simply changed the order of the polarization calculations from second-order to full-scale. Restarting the simulation resulted in a failure of the simulation after a few thousand Monte Carlo configurations due to significant growth of the magnitude of polarization energy. Similarly, changing the order of the polarization calculations from second-order to first-order and allowing the system to reach equilibrium led to the average volume and the magnitude of the average energy being underestimated by about 13% and 30%, respectively. Fortunately, the average energy and volume for simulations employing both models can be brought to near agreement with the second-order calculations with minor reparametrization.

Table 2-1. Average Total Energy, Polarization Energy (both in kcal/mol), and Volume (in Å3) for POSSIM (Second-Order Polarization), Full-Scale, and First-Order Polarizable Water Models[a]

polarization model	total energy	polarization energy	volume
full-scale	−2001 ± 35	−1411 ± 52	6335 ± 67
2nd order (POSSIM)	−1987 ± 12	−1128 ± 14	6563 ± 71
1st order	−2002 ± 19	−1100 ± 18	6665 ± 80

[a]The full-scale and first-order models use POSSIM water with refitted polarizabilities. See text for details. Uncertainties given as standard deviations of 2×10^5 configuration averages.

The full-scale and first-order results were calculated by taking our POSSIM water model and adjusting the oxygen polarizability until the average energy and volume were in best agreement with the second-order averages. While adjusting the oxygen polarizability, we maintained the ratio of oxygen polarizability to hydrogen polarizability. Decreasing the second-order polarizabilities by about 21% (from $\alpha_O/\alpha_H = 0.77/0.30$ for second-order to $\alpha_O/\alpha_H = 0.61/0.24$ for full-scale) brought the full-scale polarizability calculations within 1% and 3% of the average energy and volume respectively of the POSSIM calculations. Increasing the second-order polarizabilities by about 60% (to $\alpha_O/\alpha_H = 1.23/0.49$) brought the first-order polarizability calculations within 1% of the average energy and 2% of the average volume of the POSSIM calculations.

Interestingly, while the average polarization energy for POSSIM and the first-order calculations is nearly the same (−1128 ± 14 kcal/mol or −5.22 ± 0.06 kcal/mol per molecule for POSSIM and −1100 ± 18 or −5.09 ± 0.08 kcal/mol per molecule for first-order), the average polarization energy is noticeably larger for full-scale polarization (−1411 ± 52 kcal/mol or −6.53 ± 0.24 kcal/mol per molecule). This "extra" polarization energy is balanced by an increase in the stretching and bending energies as well as the slight overestimation of the volume.

One should keep in mind that there is no guarantee of transferability of these refit water models to other systems. These water models are to demonstrate that a simple reparametrization (only one degree of freedom in parameter space) was able to mostly eliminate the discrepancies between polarization orders of the liquid-phase properties. If a robust first-order or full-scale polarization model of water were desired, we would likely proceed via our standard parametrization methodology (i.e., start by fitting polarizabilites with three-body energies, then fit the partial charges and Lennard-Jones constants using gas-phase dimers, then fine-tune the parameters with liquid simulations). While the above examples cannot account for all the possible situations, they do serve as an illustration for the statement that the second-order polarization formalism employed in POSSIM is adequate in reproducing many-body effects, provided that a proper parametrization procedure is followed.

Chapter 3: Platinum Parametrization Project

Background

Cisplatin is well known to have an anticancer therapeutic effect via interactions with DNA in the cell nucleus (***16***,***17***) and thus it differs from many other anticancer drugs which work by interacting with various proteins There is a body of evidence that cisplatin binding to protein and protein-mediated transport are responsible for cellular resistance to this drug, (***7***,***18***) even though the exact mechanism of this process is still not clear. Moreover, the specific proteins implicated in cisplatin resistance are involved in copper binding and transport (***19***).Therefore, studying cisplatin binding to both DNA and copper transporter proteins is an important area in anticancer research. Additionally, there are other platinum(II) compounds that are similarly used as anticancer drugs (***4***).

After the cisplatin (or another Pt(II) drug) is administered, the metal ion is aquated, i.e. the Cl^- ligands are replaced by water molecules, promoted by the low cytosolic concentration of chloride. Coordination with DNA is possible by the displacement of these water molecules, and the same is likely the mechanism for binding of cisplatin to the proteins responsible for the cellular resistance to this drug. Additionally, crystal structures of protein-coordinated cisplatin in which the NH_3 ligands are absent are also known, but it is difficult to deduce the physiological relevance of these structures collected under conditions far from physiological (***8***).

Thus, Pt(II), cisplatin, and their interactions with proteins are topics of clear biological importance. Efforts have been made in the direction of development of computational tools that would allow simulations of these systems (***20***,***21***). But there are still aspects that need to be addressed and for which methods need to be developed and/or improved. The ability of the central Pt(II) ion of cisplatin and other platinum-containing drugs to readily change its coordination makes it clear that it's desirable to be able to simulate not only cisplatin as a whole, but also the central Pt(II) ion with some or all of its initial ligands replaced by water or other relevant metal-coordinating groups.

Furthermore, in the best-case scenario, seamless transitions between different coordination cases would have to occur over the course of simulations without the transitions being introduced "by hand" or by some semi-automated procedure based on preexisting structural knowledge and not on the physical flow of the simulations themselves. Quantum mechanical simulations can provide accurate results in this sense, but system size and length of such simulations (if true thermodynamic averaging in the form of molecular dynamics or Monte Carlo simulations is to be utilized) are limited. This is why empirical force fields remain the tool of choice in many biological simulations. However, existing empirical force field

models for metal ions – especially transition metal ions – are only well suited for applications in which the transition metal atoms remain fully bonded and do not change their coordination environment as they lose, acquire, or exchange ligands. In the simplest models, simulations represent metal ions by mere point charges (***22***) Obviously, such models can only be used to describe systems in which the geometry of molecular fragments surrounding the ion is fixed.

In other types of simulations, sophisticated and accurate ways of describing non-bonded interactions of metal ions with the environment are employed but positioning of the ions in respect to their ligands or coordinated protein fragments still rely on explicit harmonic or rigid constrains for the distances and angles involving the ion. Such an ion model essentially employs "pseudo-valence" bonds and angles for positioning of the metal ion. As a result, any significant changes of geometry, ion transfer or other processes involving changes in the ligand environment are not possible. Examples involving such constrains include AMBER (***23***) and LFMM/MOE (***24***,***25***). SIBFA and VALBOND are using a similar approach to directionality and positioning of ligands (***24***,***26***). Moreover, LFMM, SIBFA, VALBOND and other similar techniques utilize electronic structure information heavily (***24***). While they are empirical force fields, they compute certain non-bonded interactions with electronic structure methods. These quantum mechanical methods scale poorly with system size and makes it impossible for these techniques to reach the speed of force fields like AMBER and CHARMM (***27***).

Overall, the existing empirical or partially empirical techniques for simulating molecular interactions with cisplatin, Pt(II), and other transition metal ions are heavily dependent on preexisting knowledge of the complex's geometry and do not permit fully unconstrained simulations of the complexes or of metal–ion transfers. This is very different from the situation of purely organic complexes that can be simulated in unconstrained runs with most force fields. We are reporting multi-site models for Pt(II) and cisplatin created for simulating cisplatin and other coordinated Pt(II) systems with modified OPLS-AA and second-order polarizable POSSIM force fields. While these models do not require quantum mechanical runs during calculations and retain the fairly simple empirical formalism of the OPLS-AA and POSSIM frameworks, we have been able to reproduce a number of structural and energy results, including the energies of cisplatin and hydrated platinum(II) formation. This model can be called reactive, although only in the sense of Pt(II) coordination interactions. We hope that further applications of the model will permit the elucidation of the molecular-level mechanism of cisplatin's therapeutic action and cellular resistance to this drug and will potentially contribute to the creation of other platinum-containing anticancer compounds.

Pt(II) Ion Model

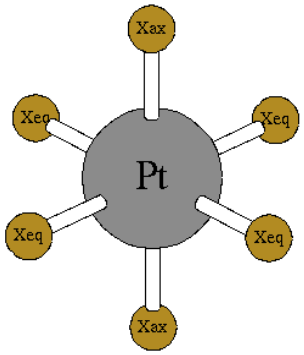


Figure 3-1. The seven-site model for the Pt(II) ion. The same model geometry was used for both the modified OPLS-AA and the POSSIM force fields. The central particle labeled "Pt" has only a Lennard-Jones potential while the virtual sites—fixed with respect to the central particle—are labeled "X" and have only Coulomb potentials. The axial virtual sites (X_{ax}) have one charge while the equatorial virtual sites (X_{eq}) have another and must satisfy $2q_{ax} + 4q_{eq} = 2$. The distance between the central atom and the virtual sites are the same in both cases. The Pt–X bonds are for visualization purposes only. The virtual sites are fixed with respect to the central atom.

With the purpose of reproducing the coordination directionality of the platinum ion interactions, the particle in Figure 3-1 was modeled by the central "atom" (Pt) with no electrostatic charge but with Lennard-Jones parameters and by six virtual sites (X_{ax} and X_{eq}) that had only electrostatic charges and no van der Waals interactions assigned to them. The bonds between the virtual sites and the central particle and all the angles were kept completely fixed. The virtual sites belonged to two different atom types (for each of the force fields). Two of the virtual sites are considered "axial" (X_{ax}) and their bonds to the central atom are separated by 180°. The other four sites are the "equatorial" (X_{eq}) type. The equatorial virtual sites are located in the plane that is perpendicular to the line between the axial virtual sites and each neighboring two sites that are separated by 90°. The spatial arrangement of the virtual sites makes the whole structure octahedral. The Pt···X distances were set to be 1.3 Å with the modified OPLS-AA and 0.87 Å with the polarizable POSSIM force field. The charges on all the equatorial sites were the same (within the same force field), but they could differ from those on the axial ones. No polarizability was assigned to this system in the POSSIM framework.

Fitting of the potential energy parameters for the platinum ion was done in the following calculations. First, the Pt(II) parameters were fit to reproduce geometries (distances) and energies of formation of the following complexes: $[Pt(H_2O)_4]^{2+}$ (Figure 3-2), $[Pt(H_2O)_6]^{2+}$, cisplatin, *cis*-

$[PtCl(NH_3)_2(H_2O)]^+$, *cis*-$[PtCl_2(NH_2CH_3)_2]$, *cis*-$[Pt(CH_3S)_2(H_2O)_2]$, *trans*-$[Pt(CH_3S)_2(H_2O)_2]$, and *cis*-$[Pt(NH_3)_2(CH_3S)(H_2O)]^+$. Then, further parameter fitting was carried out to reproduce the Pt(II) free energy of hydration in bulk water and radial distribution functions for the Pt···O(water) pair.

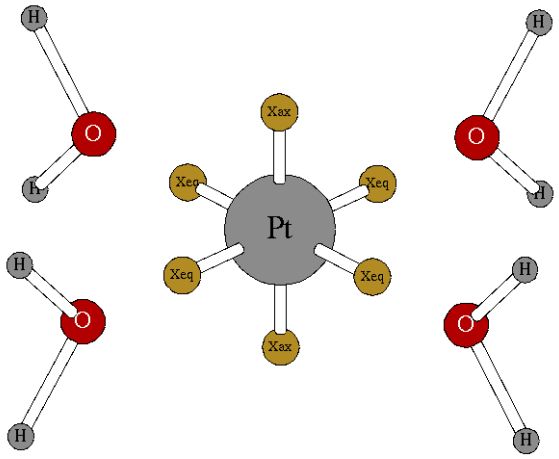


Figure 3-2. Geometry of teraaquaplatinum(II).

These, and all other quantum mechanical calculations involving platinum, were performed using the Gaussian 09 (***28***) software package at the B3PW91/mDZP/6-31+G(d) level of theory. Specifically, the calculations employed density functional theory (DFT) using the B3PW91 functional with the all-electron mDZP basis set (***29***) for platinum and the 6-31+G(d) basis set for all other atoms. We chose the level of theory and basis set using the work of Paschoal *et al.* (***29***) in which they produced an extensive survey of functional/basis set combinations (26 DFT functionals, 25 basis sets for Pt, and 36 basis sets for the ligands). The chosen functional/basis set combination was among the top three most accurate functional/basis set combinations investigated by Paschoal *et al* and has extensive literature validation in other transition metal systems. (***29***) Here, accuracy refers to accuracy in reproducing the experimental rates of hydrolysis and geometry of cisplatin. All the geometry optimizations were run with the POSSIM software suite (***11***,***15***,***30***,***31***).

Additional Parametrization

The cisplatin model required us to fit parameters for the Cl^-, NH_3, and CH_3S^- ligands in addition to the central Pt(II) ion. The parameters for the ligands were fit with the POSSIM and modified OPLS-AA formalisms. The values of the parameters were determined by carrying out gas-phase and liquid-state

simulations. Three-body energies were used to fit the polarizabilities in the standard fashion for the POSSIM parametrization We fit the Lennard-Jones parameters by using extrapolated (***14***) quantum mechanical dimerization energies (see Chapter 2 for more details). Shown in Figures 3-3 and 3-4 are the dimers used in the fitting procedure. Finally, Monte Carlo liquid-state simulations yielded heats of vaporization, densities, and hydration energy values that we compared with experimental data.

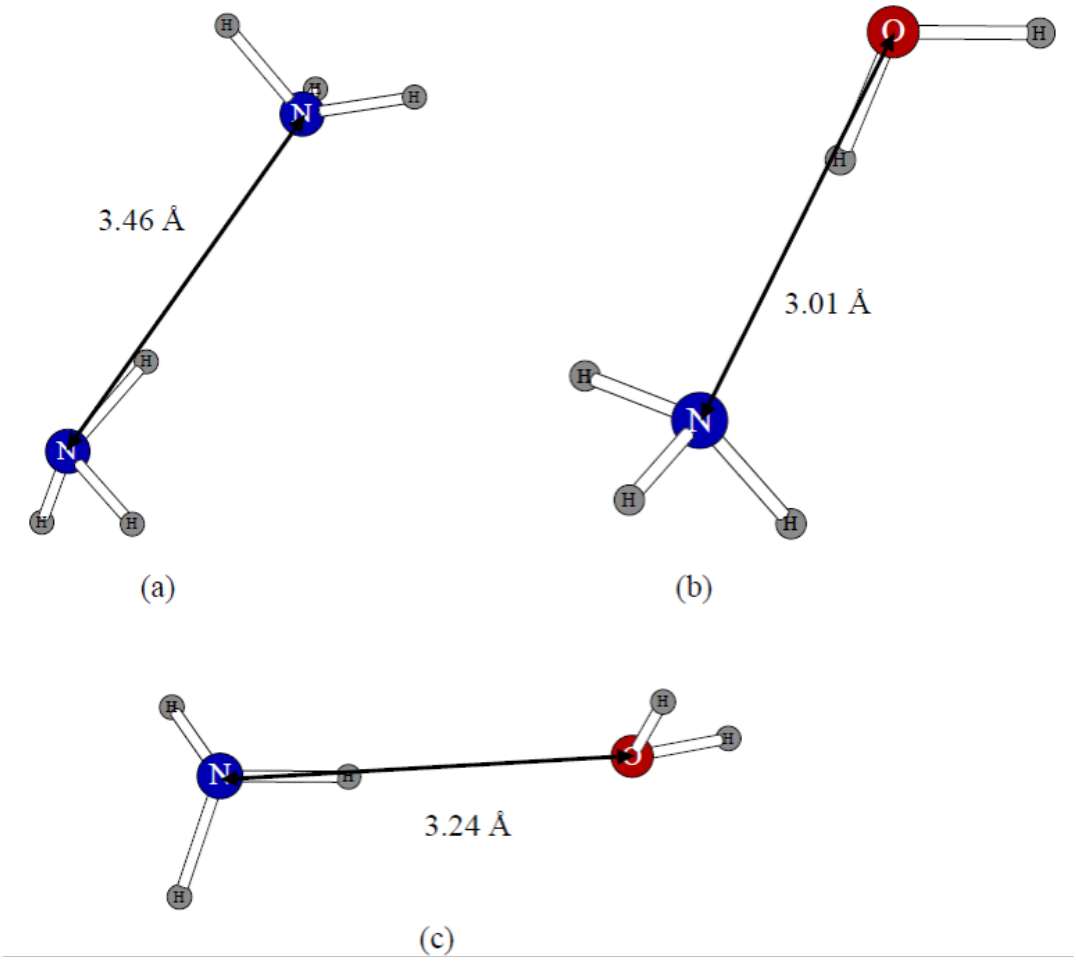


Figure 3-3. Dimers used in fitting NH_3 parameters: ammonia homodimer (a); ammonia – water dimer with N as the acceptor (b); ammonia – water dimer with O as the acceptor (c). The given distances are the minimum energy quantum mechanical values.

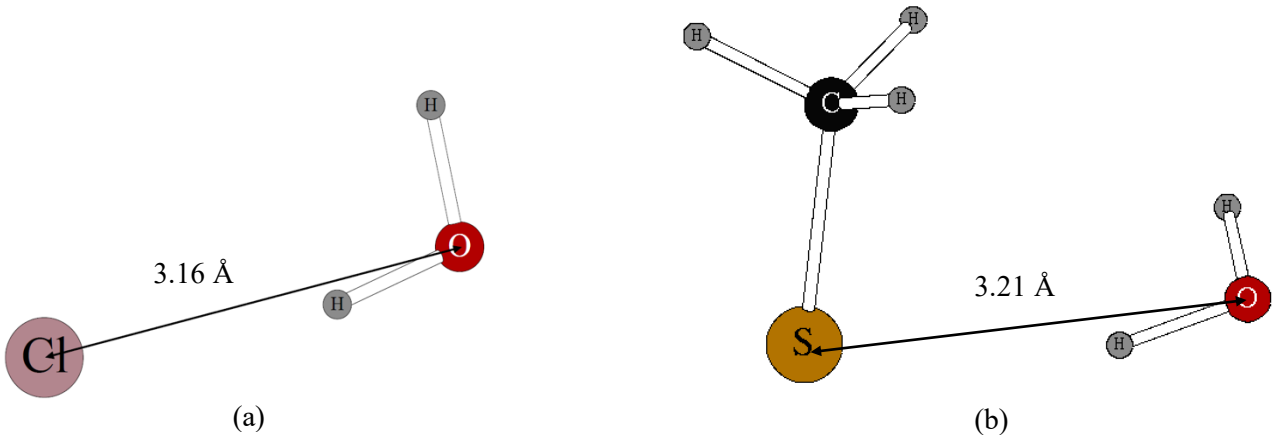


Figure 3-4. Dimers used for fitting chloride and methanethiolate parameters: chloride–water dimer (a) and methanethiolate–water dimer (b). The distances shown are from quantum mechanical energy minimization.

Pure Liquid Simulations

Simulations of neat liquid ammonia were carried out with the POSSIM 2.0 software package. These were used to fine-tune the force field parameters by reproducing experimental values of enthalpy of vaporization and density. Simulations consisted of an NPT ensemble (constant number of molecules, pressure, and temperature) of 267 molecules in a cubic box with periodic boundary conditions at 1 atm and -33.33 °C (the boiling point of ammonia). Simulations employed the Monte Carlo method with Metropolis sampling criteria. Non-bonded interactions greater than defined cutoff values were set to zero. For dipole-dipole interactions the cutoff value was 7.0 Å, the cutoff value for all other non-bonded interactions was 9.0 Å. The charge-charge interactions were quadratically feathered to zero over the last 0.5 Å before the cutoff distance. This technique reduces unnecessary noise associated with charges moving into and out of the cutoff radius. The following correction to the Lennar-Jones energy is applied to correct for ignoring interactions beyond the cutoff distance:

$$E_{LJ\,cor} = \frac{\rho}{2}\int_{R_{cut}}^{\infty} V_{LJ}(r)\,4\pi r^2 dr \qquad (17)$$

where ρ is the molar density, $V_{LJ}(r)$ is the Lennar-Jones potential, and R_{cut} is the cutoff value.

Each simulation comprised at least 2×10^6 configurations of equilibration followed by at least 4×10^6 configurations of averaging. Each simulation was broken up into batches of 2×10^5 steps and the variance calculated by the batch means method:

$$\sigma^2 = \sum_i^m \frac{(X_i - \bar{X})^2}{m(m-1)} \tag{18}$$

where m is the number of batches, X_i is the average of some property of the ith batch. Uncertainties are reported as the standard deviation (± σ). The enthalpy of vaporization is calculated as follows:

$$\Delta H_{vap} = E(gas) - E(liq) + RT \tag{19}$$

Where *E(gas)* and *E(liq)* are the average energy of the gas-phase and liquid respectively, R the gas constant, T the temperature. The RT term is included to account for the Δ(PV) part of the enthalpy under the assumption that dilute ammonia behaves as an ideal gas and the molecular volume of the liquid is negligible compared to that of the gas. The density is calculated directly from the average molar volume and the molecular mass of ammonia.

Calculating Free Energies of Hydration

Free energies of hydration were calculated as a test of the fitted parameters and were not used in the fitting procedure. In contrast to the simulations employing periodic boundary conditions for pure ammonia, simulations to calculate the free energies of hydrations were carried out in water droplets. These droplets consisted of a solute, fixed at the center of the simulation box, surrounded by 216 water molecules. In order to prevent water molecules on the surface of the cluster from "evaporating," a half-harmonic restoring force of 1 kcal/mol/Å^2, originating at the center of the box, was applied. All simulations were carried out at 25°C with the POSSIM 2.0 software package.

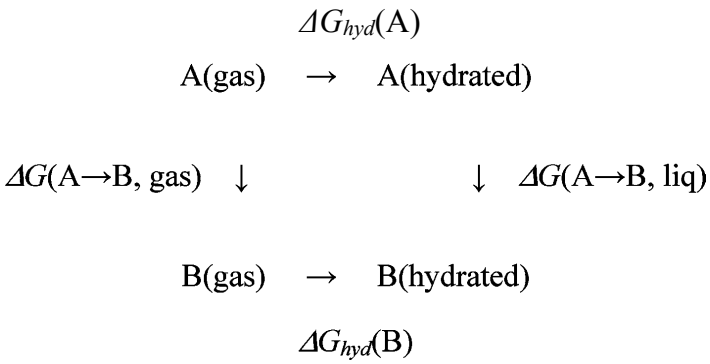


Figure 3-5. Thermodynamic cycle used to calculate free energies of hydration.

The thermodynamic cycle used to calculate the relative free energies of hydration between species A and B is given in Figure 3-5. From this cycle, the relative free energy of hydration between A and B is:

$$\begin{aligned}\Delta\Delta G_{hyd} &= \Delta G_{hyd}(B) - \Delta G_{hyd}(A) \\ &= \Delta G(A \rightarrow B, liq) - \Delta G(A \rightarrow B, gas)\end{aligned} \tag{20}$$

However, if species B is a dummy atom (an atom with no intramolecular interactions), $\Delta G_{hyd}(B)$ equals zero and Equation 20 affords the absolute free energy of hydration of A. In this case, the two $\Delta G(A \rightarrow B)$ terms can be thought of as the free energy of "annihilating" species A. These free energies were calculated using statistical perturbation theory:

$$\Delta G(i \rightarrow j) = -RT \ln\langle e^{-(E_j - E_i)/RT}\rangle_i \tag{21}$$

where the brackets, $\langle\ \rangle_i$, indicate averaging the value inside the brackets over the configurational space of the system at point i while E_i and E_j are the energy at points i and j respectively. However, since the change between a species A and a dummy atom is too large to accurately sample, the process is broken up into stages. First the polarizability is removed (*i.e.* A is changed to a species that is the same as A except it lacks polarizability), then the charges are removed, and finally, the Lennard-Jones parameters are removed. Each of these stages is further broken down into a series of steps with the parameter decreasing stepwise from its value in A to zero.

Results and Discussion

Ammonia

For the non-polarizable model, the parameters for ammonia were taken from the published OPLS parameter set for amines (***32***) and the results reproduced using the POSSIM software package. For the polarizable model, the average deviation of the three-body energies was 0.002 kcal/mol.

Table 3-1 compares the energy-minimized dimer structures and energies calculated with the polarizable model (POSSIM) to the calculated quantum mechanical reference values and the published OPLS results. The geometry reproduced by the POSSIM polarizable force field is slightly better than the geometry reproduced by OPLS. The average deviation of the distance from the quantum mechanical value reproduced by the POSSIM polarizable field is 0.21 Å, while the OPLS deviation is 0.24 Å. While the energy error is smaller with the OPLS formalism, fixed-charges force fields are known to overestimate gas-phase binding energy, and the quantum mechanical result of –6.55 kcal/mol for the second dimer may be overestimated, especially in the view of similar trends we observed with parametrizing nitrogen containing compounds in the past (***11***).

Table 3-1. Quantum mechanical, POSSIM, and OPLS equilibrium heavy atom···heavy atom distances and dimerization energies for: ammonia as H-bond donor (to water), ammonia as H-bond acceptor (from water), and the ammonia homodimer.

System	N···O/N **distance,** Å			**Dimerization E,** kcal/mol		
	QM	POSSIM	OPLS	QM	POSSIM	OPLS
O as H-bond acceptor	3.24	3.13	3.03	–2.58	–2.32	–3.10
N as H-bond acceptor	3.01	2.87	2.85	–6.55	–5.09[a]	–6.50
NH_3···NH_3 dimer	3.46	3.14	3.15	–2.96	–2.39	–3.11
RMS deviation from QM		**0.21**	**0.24**		**0.44**	**0.31**

[a]Not included in RMS deviation

Liquid state results are given in Table 3-2. Both POSSIM and OPLS-AA produce accurate values of heats of vaporization and densities within 0.5-3% of the experimental data, though the second-order polarizable POSSIM formalism performs slightly better.

Table 3-2. Comparison of experimental and calculated enthalpies of vaporization and densities for liquid ammonia at its boiling point (–33.33 °C) at standard pressure.

Method	**ΔH_{vap},** kcal/mol	**deviation from experiment**	**density,** g/cm^3	**deviation from experiment**
experimental	5.58[a]		0.682[a]	
POSSIM	5.56 ± 0.01	-0.4%	0.688 ± 0.003	0.9%
OPLS	5.42 ± 0.01[b]	-2.9%	0.697 ± 0.001[b]	2.2%

[a]Data from (***33***) [b]Data from (***34***)

Chloride

In order to parametrize the OPLS model, we surveyed three published chloride parameter sets and selected the best set for refinement. Significant improvements to these parameters could not be found, so the parameters (***35***) were used without modification and the published values reproduced in the POSSIM software package. However, parametrizing the POSSIM model of the Cl^- ion using the B3LYP/cc-pVTZ(-f) level of theory resulted in three-body energies and polarizabilities that were qualitatively much too small for a halide. Therefore, we used a published value for an estimate of the polarizability of chloride in bulk water (***36***), which proposed $\alpha = 5.0$ Å^3 which is significantly larger than previous estimates (***37***) (~4 Å^3) and only slightly reduced from the gas-phase value (***38***) (5.48 Å^3). Table 3-3 contains the results of fitting POSSIM parameters for this ion, as well as results employing the published OPLS parameters. The results demonstrate that both POSSIM and OPLS perform adequately in the gas-phase dimerization calculations, with POSSIM being somewhat more accurate than OPLS in structure.

Calculating the free energy of hydration with the statistical perturbation energy leads to the fixed-charges OPLS simulations yielding an error of 2.4% while the POSSIM simulations were significantly more accurate with an error of only 0.3%. The TIP3P water model (***39***) was used in the OPLS-type simulations and our POSSIM water model was used for the polarizable calculations

Table 3-3. Gas-phase Cl^-···water dimer and Cl^- hydration free energy in bulk water.

Method	**Dimerization E**, kcal/mol	Cl···O **distance**, Å	**ΔG_{hyd},** kcal/mol
Experimental	–14.5 ± 0.5[a]		–72.7 ± 2[b]
QM	–13.65	3.16	
POSSIM	–15.08	3.16	–72.5 ± 1.1
OPLS-TIP3P	–13.61	3.19	–74.4 ± 1.7

[a]Data from (***40***) [b]Data from (***41***)

Methanethiolate

Parameters for methanethiol (the side chain analog for protonated cystine), were developed previously (***15***). However, further fitting was need for CH_3S^- (the side chain analog for deprotonated cysteine) because of the central role of the deprotonated cysteine residue in coordinating transition metal ions including the platinum(II) ion of cisplatin.

In order to reflect the longer H-bond distances of sulfur compounds, we needed to modify our standard procedure for calculating three-body energies (as described in ref (***12***) and Chapter 2) for methanethiolate. For sulfur-containing species, we increase the distance from the S atom to the dipolar probes to 2.1 Å from the usual 1.8 Å. The average unsigned error in the three-body energy of methanethiolate was 0.031 kcal/mol. These results are consistent with the usual level of accuracy we achieve in reproducing three-body energies (***15***). The results of fitting POSSIM and OPLS parameters for methanethiolate are shown in Table 3-4. Both methods were capable of reproducing the gas-phase dimerization energy (to within the bound of the QM and experimental references), the S···O dimer distance (to within 0.05 Å), as well as the bulk free energy of hydration (to within the experimental uncertainty).

Table 3-4. Gas-phase CH_3S^-···water dimer and CH_3S^- hydration free energy in bulk water.

Method	**Dimerization E,** kcal/mol	S···O **distance,** Å	**ΔG_{hyd},** kcal/mol
Experimental	–15.0 ± 0.2[a]		–73.8 ± 2[b]
QM	–13.28	3.21	
POSSIM	–14.70	3.24	–73.9 ± 1.7
OPLS-TIP3P	–14.14	3.16	–75.6 ± 1.9

[a]Data from (***42***) [b]Data from (***43***)

Overall, both POSSIM and OPLS-AA parameters for ammonia, the chloride ion, and the methanethiolate ion appear to robust enough for our purposes and were used in simulations of cisplatin and platinum-containing complexes without additional refitting.

Platinum(II) and Complexes of Platinum(II)

Consider the simplest platinum complex we investigated—$[Pt(H_2O)_4]^{2+}$, depicted with the seven-site platinum model in Figure 3-2. As a fitting target, we used the QM-calculated geometry and energy of formation of this system. Here, the energy of formation is the zero-point corrected enthalpy of reaction (at 0 K) of the molecular fragments (the Pt^{2+} ion and four water molecules) to form the $[Pt(H_2O)_4]^{2+}$ complex. This treatment was applied to all of the complexes investigated. Tables 3-5 and 3-6 present our results (for both POSSIM and OPLS) for fitting platinum force field parameters with $[Pt(H_2O)_4]^{2+}$ and the other platinum complexes used as fitting targets in comparison with the QM-calculated values. Table 3-5 contains the energies of formation of the fitting targets while Table 3-6 contains key bond lengths.

Table 3-5. Total energy of formation (kcal/mol) for cisplatin, cisplatin hydrolysis products, and thiolate complexes.

	Method		
	QM	POSSIM	OPLS
$[Pt(H_2O)_4]^{2+}$	–297.8	–294.0	–361.2
$[Pt(H_2O)_6]^{2+}$ [a]	–327.8	–331.6	–400.5
cisplatin	–680.3	–638.0	–613.0
cis-$[PtCl(NH_3)_2(H_2O)]^+$	–568.0	–544.5	–540.2
cis-$[PtCl_2(NH_2CH_3)_2]$	–683.5	–646.9	–604.3
cis-$[Pt(CH_3S)_{2)}(H_2O)_2]$	–703.6	–604.6	–622.2
trans-$[Pt(CH_3S)_2(H_2O)_2]$	–695.5	–612.8	–634.5
cis-$[Pt(NH_3)_2(CH_3S)(H_2O)]^+$	–595.5	–530.1	–522.5
RMS deviation from QM		**55.4**	**67.6**

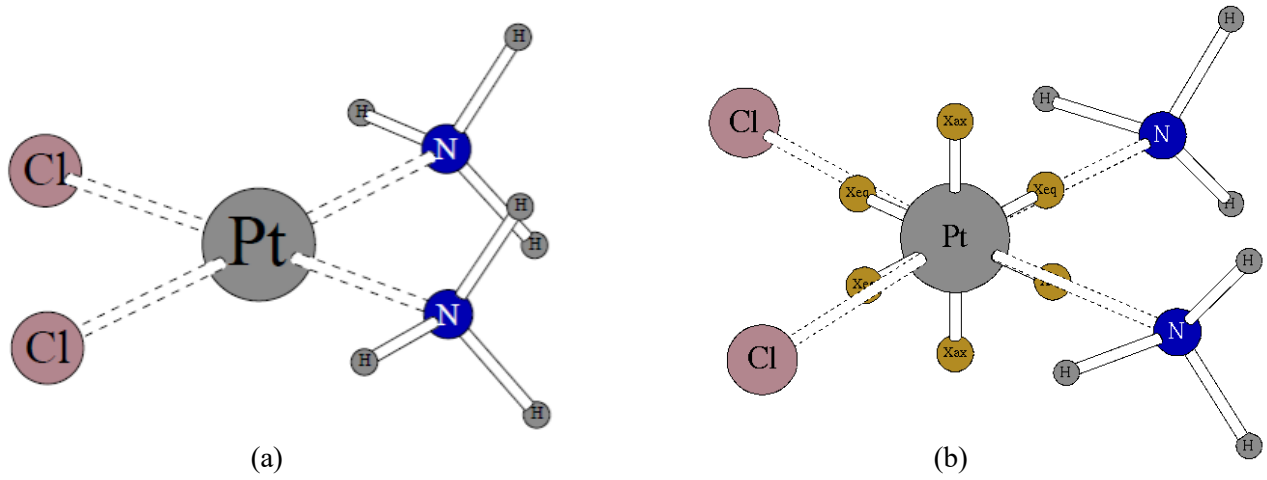
[a] Not included in RMS deviation

The results demonstrate that the fixed-charges modified OPLS and the polarizable POSSIM force fields produce geometries in a good agreement with quantum mechanical data (deviations in distances are within ca. 0.1Å). At the same time, the POSSIM formalism permits the energy of formation to be much closer than its quantum mechanical counterpart (–294.0 vs. –297.8 kcal/mol). Note that these calculations of the geometry and energy of formation of the coordination complex have been carried out with empirical force field that contains no additional special energy terms.

Table 3-6. Structural details of cisplatin and hydrolysis products and thiolate complexes

	Distance, Å								
	QM			POSSIM			OPLS		
	Pt···Cl/S	Pt···N	Pt···O	Pt···Cl/S	Pt···N	Pt···O	Pt···Cl/S	Pt···N	Pt···O
$[Pt(H_2O)_4]^{2+}$			2.06			2.11			2.02
$[Pt(H_2O)_6]^{2+}$ [a]			2.06/2.98			2.14/2.50			2.03/2.63
cisplatin	2.34	2.11		2.39	2.25		2.45	2.15	
cis-$[PtCl(NH_3)_2(H_2O)]^+$	2.31	2.09	2.31	2.27	2.14	2.34	2.42	2.14	2.03
cis-$[PtCl_2(NH_2CH_3)]$	2.35	2.11		2.39	2.19		2.45	2.11	
cis-$[Pt(SCH_3)_{2)}(H_2O)_2]$	2.32		2.21	2.42		2.21	2.42		2.03
trans-$[Pt(SCH_3)_2(H_2O)_2]$	2.38		2.10	2.40		2.24	2.41		2.03
cis-$[Pt(NH_3)_2(SCH_3)(H_2O)]^+$	2.33		2.12	2.35	2.09	2.22	2.40	2.14	2.03
RMS deviation from QM					**0.09**			**0.08**	

[a] not included in RMS deviation



(a) (b)

Figure 3-6. Calculated geometry of cisplatin obtained with quantum mechanical (a) and POSSIM (b) simulations

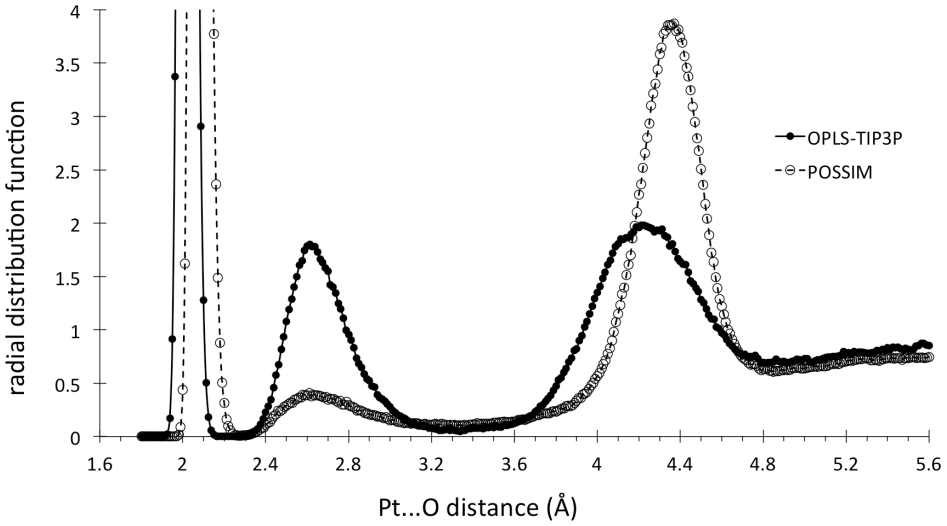


Figure 3-7. Pt···O radial distribution function. For POSSIM, the first and second maxima are at 2.07 Å and 2.62 Å. For OPLS with TIP3P water, the first and second maxima are at 2.02 Å and 2.65 Å. The experimental results are 2.02 Å and 2.74 Å.

All of the results show that the agreement in interatomic distances between the empirical and quantum results is consistently good with an RMS deviation of 0.08Å for modified OPLS and 0.09Å for POSSIM. The formation energies are simulated somewhat more accurately with the polarized POSSIM force field with an RMS deviation from QM of 55 kcal/mol compared to 67 kcal/mol for modified OPLS. In addition to validation against the above quantum mechanical gas-phase complex formation results, we carried out simulations aimed at comparison of our simulation performance with experimental data on the Pt(II) ion solvation. The results are summarized in Table 3-7 and the associated radial distribution functions (RDF) shown in Figure 3-7. The locations of the first and second RDF maxima are within ca. 0.1Å of the experimental data, with the modified OPLS giving a slight advantage. The free energy of hydration is simulated more accurately with POSSIM, as the difference with the experimental data is only 0.9%, essentially within the error bar, while the modified OPLS deviation is about 5%. Therefore, once again, the both models work adequately enough, but the energy-related properties are calculated more accurately when the polarizable POSSIM methodology is employed.

Table 3-7. Free energy of hydration of Pt(II) and locations of first two maxima of the $Pt^{2+}\cdots OH_2$ radial distribution function.

	ΔG_{hyd}, kcal/mol	Pt···O **distance**, Å	
Experimental	–483.8 ± 4 [a]	2.02 [b]	2.74 [b]
POSSIM	–488.3 ± 1.8	2.07	2.62
OPLS-TIP3P	–509.2 ± 2.0	2.02	2.65

[a] Calculated in the manner of (***41***) with data from (***44***) [b] Data from (***45***)

Overall, we have achieved a robust level of accuracy in obtaining the energy and structure of Pt(II) coordination compounds and Pt(II) hydration by using empirical force fields. Our model demonstrates transferability from the model compounds used in fitting to more complicated systems of biological interest. Furthermore, since bond, bond angle, and dihedral angle terms are not used to describe the metal–ligand interactions, specific Pt–ligand parameters are not needed and thus preexisting knowledge of complex geometries is not required. The polarizable POSSIM force field gives an advantage when energy properties are to be calculated. Importantly, our implementation of the multisite model for the platinum ion permitted us to reproduce the experimentally and quantum mechanically observed nearly-squared planar coordination geometry, as can be seen from the structures presented in Figure 3-6.

Given the results of our project, we expect that both sets of parameters (modified OPLS-AA and POSSIM) should work reasonably well in reproducing geometries of other Pt(II) coordination complexes (such as complexes of this ion with proteins and DNA). However, the polarizable POSSIM version outperforms the fixed-charges modified OPLS-AA in reproducing and predicting energies of complex formation. The values for all force field parameters derived and used in this work can be found in Table A-1

Conclusions

We have created and tested polarizable (POSSIM) and fixed-charges (modified OPLS-AA) models for Pt(II) and Pt(II) coordination compounds. The investigated coordination compounds include cisplatin and its key hydrated variants as well as platinum—thiolate complexes. We parametrized these models by fitting to high-level quantum mechanical energies of formation—calculations which cannot be performed

accurately with traditional force field approaches. Standard force field models of transition metals use pseudo-bonds (and often harmonic stretching and periodic torsional terms as well) connecting the ion with the coordinated ligands. This makes the relative positions of the ions and ligands predetermined. Furthermore, the coordination of the metal ion cannot change during the course of a simulation unless parameters are deliberately changed based on information external to the force field (e.g. investigator chemical intuition or some employed heuristic.

In contrast, the main distinguishing feature of our models is that structure and energy is simulated without knowledge of the specific system study or geometric constraints. Rather, the metal—interactions in our models are based on the nonbonded interactions of a simple empirical force field. Finally, the coordination of the Pt(II) ion can change during the course of a simulation without any additional actions beyond the normal application of the force field.

Chapter 4: Binding of Copper and Cisplatin to Atox1

Introduction

Copper is an essential micronutrient required for many key biological processes, including oxidative phosphorylation (***46***), hormone and melanin production (***47***), neurotransmission (***48***), protection against reactive oxygen species (***49***), and many others (***50***). Deregulation of copper distribution and metabolism plays a role in the development of many debilitating and progressive diseases, such as Wilson disease (***50***), Menkes disease (***50***,***51***), Alzheimer's (***51***,***52***), Parkinson's (***52***), and cancer (***53***). Due to its ability to cause oxidative damage through production of reactive oxygen species, copper is not found inside the cell in the free state, but it is always bound to the proteins and low molecular weight compounds such as glutathione (***54***).

Copper enters the cell through a passive transporter *h*Ctr1 located in the plasma membrane. Atox1, a small ferredoxin-like chaperone protein, accepts copper from *h*Ctr1 and delivers it to ATP-driven transmembrane copper transporters ATP7B and ATP7A for incorporation into the proteins or removal of excess copper from the cell. All of these copper transport proteins have been shown to be involved in cell resistance to cisplatin (***7***,***55***), a platinum-based chemotherapy agent, effective against many solid tumors including ovarian, testicular, cervical cancer, melanoma and others (***56***). Cisplatin binds to the highly conserved copper binding CxxC motif in ATP7A/B (***19***,***57***), and Atox1 (***8***), as well as to the methionine-rich motifs presumed to be involved in copper transport in *h*Ctr1 (***58***,***59***). Furthermore, in vitro experiments have shown that the copper chaperone Atox1 can transfer cisplatin to the metal-binding domain of ATP7B (***19***) and ATP7A (***60***). Intriguingly, experiments performed using a model vesicle system suggest that ATP7B may facilitate the transport of cisplatin across cellular membranes (***61***). Finally, the deletion of Atox1 in both *Drosophila melanogaster* (***62***) and mouse fibroblasts (***63***) leads to reduced sensitivity to cisplatin. These observations suggest that cisplatin can be transported through the cell along the copper transfer pathways.

DNA is considered to be the main pharmacological target of cisplatin: the drug binds to guanine, causing intra- and inter-strand cross-links that impair DNA transcription and replication, causing apoptosis (***16***,***17***). Interestingly, Atox1 has been shown to deliver copper into the nucleus (***64***) and act as a transcription factor (***65***). This suggests that Atox1 could also deliver cisplatin to the nucleus, which could have important implications in research and development of platinum-based chemotherapeutics (***19***).

If platinum and copper bind to the same site in Atox1, one might expect a competitive relationship between the two metals. However, Atox1 can bind both copper and cisplatin simultaneously (***60***,***66***,***67***)

Moreover, in the presence of glutathione cisplatin was shown to react with Atox1–Cu faster than with *apo*-Atox1 (***60***). Binding of copper and cisplatin to Atox1 in the experiments published so far was carried out either in anaerobic conditions without the presence of any reducing agents, or in the presence of low concentrations of dithiothreitol (DTT). The goal our collaboration was to analyze binding of copper and cisplatin to Atox1 under physiological redox conditions, and to determine the chemical structure of the Atox1–Cu–Pt complex.

Our experimental collaborators simulated the in-cell redox environment by using partially oxidized glutathione with the redox potential of the reduced/oxidized glutathione pair (GSH/GSSG) within the range found in the cytosol of mammalian cells (***68***). Under these conditions, they found that Atox1–Cu reacts with cisplatin, forming complexes in which copper and platinum are bound to the protein through metal–sulfur clusters, which include glutathione. Detailed experimental methods can be found in Reference (***69***); key experimental evidence is discussed below, in the "Results and Discussion" section.

Computational Methods

Structure Preparation

In order to expedite molecular mechanics calculations, we modeled Atox1 as residues 9 to 19 of the Cu(I)-loaded crystal structure (PBD ID: 1FEE) (***70***) with the addition of C- and N-terminal capping groups. Use of this truncated model is justified as it includes all the residues showing significant change in NMR chemical shift upon binding of Cu(I) and cisplatin to *apo*-Atox1 in solution, with the exception of K60. In addition, the solution structure (PDB ID: 1TL4) (***71***) and crystal structure (PBD ID: 1FEE) of Cu(I)-loaded Atox1 show no major structural differences with an RMSD of ca. 0.9 Å—comparable to the RMSD of 0.95 Å for the solution-phase NMR ensembles (***71***).

Hydrogen atoms were added, and the protonation states of ionizable residues were set to their most stable forms at pH 7 as applicable using the pepz utility of the MCPRO software package (***72***). Subsequently, the cysteine residues involved in copper binding were set to their deprotonated thiolate forms. The pepz utility was also used to build the naïve starting geometry of glutathione as a linear peptide with idealized bond lengths and angles.

Monte Carlo Molecular Mechanics Calculations

We performed all nonpolarizable calculations with the MCPRO software package within the OPLS force field framework (***72***). The standard OPLS parameter set (***9***,***72***) was used for all atoms except for the

thiolate sidechain of cysteine, Cu(I), and Pt(II) which we have previously parametrized (***31***,***73***,***74***). The TIP3P water model (***39***) was used in all aqueous-phase simulations.

We built protein complexes and screened for stability using a defined methodology. First, we assembled the desired complex from the truncated Atox1 model, glutathione, and metals ions. Then we performed a gas-phase optimization with only the intermolecular and cysteine sidechain degrees of freedom free. Next, we solvated the complex in an appropriately sized water droplet with a 1 kcal/mol/Å^2 half harmonic restoring potential with equilibrium distance equal to the radius of the droplet. With all solute degrees of freedom fixed, we allowed the water droplet to equilibrate for ca. 1×10^6 MC steps. Next, we applied large (>150 kcal/mol/Å^2) harmonic potentials between metal and coordinated sulfur atoms with equilibrium distances equal to the optimized gas-phase distances, and the system was allowed to equilibrate for at least an additional 1×10^6 steps. During this "relaxation" step, all intermolecular, internal glutathione, and Atox1 cysteine sidechain degrees of freedom were allowed to vary. Finally, we removed the metal ligand harmonic potentials and performed at least an additional 2×10^6 MC steps while monitoring key metal–sulfur distances. We deemed complexes in which one or more of these distances diverged, indicating complex dissociation or fragmentation, unstable and discarded. We subjected stable complexes to an additional 2×10^6 steps of equilibration followed by 8×10^6 configurations of averaging. Each averaging run was broken up into batches of 2×10^5 steps and the variance calculated by the batch means method.

We performed polarizable calculations with our POSSIM software suite utilizing the second order polarizable POSSIM force field (***30***,***15***) We used the standard POSSIM parameters (***15***) and our polarizable POSSIM water model (***11***), except for the thiolate sidechain of cysteine and Pt(II) which we have recently parametrized (***74***). We obtained parameters for the Cu(I) ion in the same manner as for our previous generation polarizable force field (***31***). Due to the combined size of these protein complexes and solvent (>1500 atoms), we only performed polarizable calculations on structures that had passed the nonpolarizable screening process. We used the equilibrated nonpolarizable complexes and water droplet as starting geometries after truncating the glutathione molecules to methyl-capped cysteine residues and removing extraneous water molecules.

Density functional Theory Simulations

Our theoretical collaborators performed DFT-level calculations by using a hybrid orbital-free/Kohn-Sham DFT method (***75***) which enables the use of explicit solvent in DFT simulations. In this

approach, the chemically active parts of the system, including the first solvation shells, are treated at a Kohn-Sham (KS) DFT level, while an approximate orbital-free (OF) DFT is used for the remaining solvent molecules. The implementation of the hybrid OF/KS DFT approach is based on the RMG code, in which Kohn-Sham equations are solved in real-space and the multigrid technique is used to accelerate convergence of the ground state wavefunctions (***76***,***77***). For their grid-based calculations, our collaborators used a grid with spacing of 0.32 Bohr and a corresponding kinetic energy cutoff of 50 Ry. Additionally, they employed ultrasoft pseudopotentials and a generalized gradient approximation in the PBE form (***78***).

Our collaborators used our pre-equilibrated MC/MM structures as the starting point for their DFT calculations Additionally, they applied a similar truncation to both Atox1 and glutathione for increased computational efficiency. Specifically, Atox1 peptides were represented by Cys-Gly-Gly-Cys fragments with backbone extended on each side until the next α-carbon which were capped with CH_3 groups. Our collaborators applied the same truncation to glutathione (the tripeptide γ-Glu-Cys-Gly) but at the α-carbons of Glu and Gly. These fragments were fully solvated and structurally optimized within the OF/KS DFT method.

Results and Discussion

Chemical structure of Atox1–Cu–Pt complex

As our goal was to investigate the binding of copper and cisplatin to Atox1 under physiological redox conditions, reactions of Atox1 with equimolar concentrations of copper and cisplatin were carried out in the presence of 9.5 mM reduced glutathione (GSH) and 0.5 mM of oxidized glutathione (GSSG) at pH 7.4. The potential of the GSH/GSSG redox pair under these conditions is approximately –230 mV which is within the range reported for proliferating mammalian cells (***79***).

From our collaborators' experimental data, we conclude: (i) In the presence of glutathione and Cu(I), Atox1 exists as a dimer. (ii) After the addition of cisplatin, an Atox1–Cu–Pt complex is formed, in which copper is bound to three S atoms and platinum is bound to four S atoms. (iii) Within the metal0binding region of Atox1, Cu and Pt bind in close proximity, but do not form a metal–metal bond.

We have considered possible chemical structures that satisfy these conditions, and we tested several models by molecular mechanics and DFT calculations. The copper EXAFS Fourier transform peaks at around 2.7 Å are characteristic of the Cu···Cu interactions in Cu_4S_6 clusters found in copper loaded metallothioneins (***80***,***81***) (Figure 4-1 F). We based our initial models on metal arrangements in the

Cu_4S_6 clusters because the Cu–glutathione complexes and Atox1–Cu–Pt show very similar 2.7 Å peaks in their spectra.

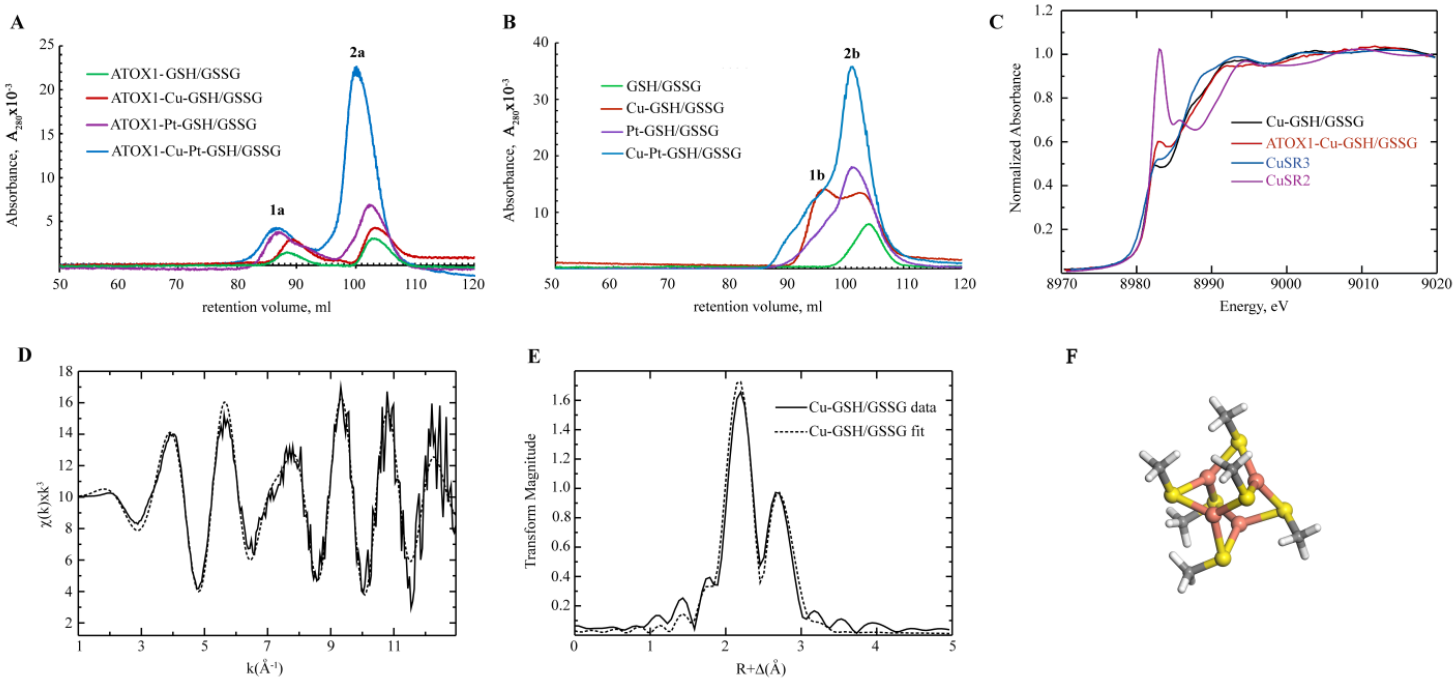


Figure 4-1. Characterization of Atox1 and glutathione complexes with copper and cisplatin. **A:** Size exclusion chromatography (SEC) profiles of Atox1 with GSH/GSSH with and without copper (Cu) and/or cisplatin (Pt). **B:** SEC chromatography profiles of GSH/GSSH with and without copper and/or cisplatin. **C:** Copper K-edge spectra of copper(I)–glutathione polymers (black) and Atox1–Cu in the presence of glutathione (red), shown together with model Cu(I) compounds for trigonal (blue, $Cu(SR)_3$) and diagonal (purple, $Cu(SR)_2$) geometries. **D, E:** Copper EXAFS spectrum (k^3-weighted) (**D**) and the corresponding Fourier transform (**E**) of copper(I)–glutathione polymers. Experimental data is shown as solid line, fitted data as dashed line. **F:** A model of Cu_4S_6 cluster with copper and sulfur atoms displayed in orange and yellow, respectively.

To create a benchmark, we used Monte Carlo molecular mechanics using both the fixed-charge OPLS and polarizable POSSIM force fields to generate hydrated structures of an 11-residue fragment of Atox1 monomer containing the CGGC metal-binding motif and flanking residues, with a bound Cu(I) atom. Both models reproduced the copper bound geometry with similar levels of accuracy. The average Cu–S distances were 2.21(5) Å and 2.28(7) Å for the OPLS and POSSIM models respectively. Both models exhibited nearly linear (ca. 150°) S–Cu–S coordination. This compares favorably to previous NMR and EXAFS studies of the Atox1–Cu monomer that indicated diagonal copper coordination geometry with a Cu–S distance of 2.17 Å and S–Cu–S angle of about 160° (***82***,***71***). However, neither

model was able to reproduce our colleagues' experimental data that consistently showed the formation of tricoordinated copper (Cu–S distances of ~2.25 Å) in systems containing Atox1, copper(I), and glutathione. Interestingly, neither the non-polarizable nor the polarizable model resulted in a stable coordination structure as the S^- of glutathione consistently migrated to the second solvation shell (~3.6 Å) of the copper atom. Attempts to produce dimeric structures of copper-loaded Atox1 with a bridging glutathione ligand were also unsuccessful as these structures consistently disassociated into copper-loaded monomers.

This strong preference for dicoordinated structures in disagreement with experimental data, which indicates CuS_3 coordination mode, is best explained by the limitations of the truncated Atox1 model used in our simulations. Specifically, the model does not include Lys60, which is proximal to the metal binding domain (Figure 4-2 D), and has been shown to be necessary for optimal copper transfer to ATP7B (***83***). Presumably, the positively charged lysine sidechain stabilizes the net negative charge of the copper–thiolate complex. Atox1 dimerization may also provide additional stabilization for tricoordinate copper coordination by sulfur observed experimentally.

We then screened a number of potential Atox1–Cu–Pt–glutathione structures for stability. In order to reduce the number of possible structures tested, two key factors were considered. First was agreement with the number and type of metal-coordinated ligands as shown by our colleagues' spectroscopic data. Second, as indicated by the NMR data shown in Figure 4-2, only structures in which copper is bound to one or more of the cysteine residues in Atox1 dimer were considered. Due to the sub-femtomolar binding affinity of Atox1 for Cu(I) (***84***), Pt(II) would not be expected to completely displace copper from the binding site.

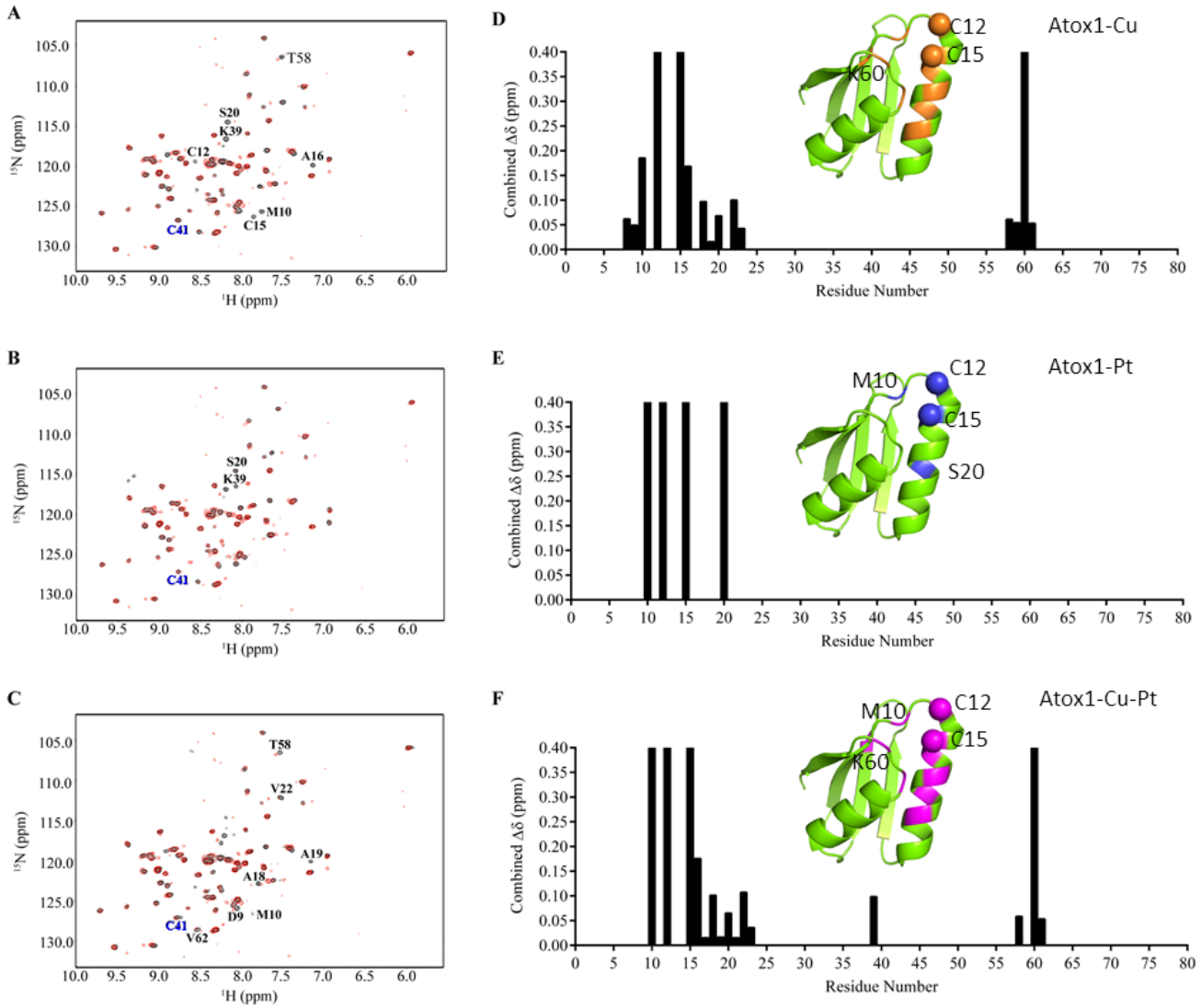


Figure 4-2. Binding of cisplatin to Atox1–Cu in the presence of glutathione. **A-C:** Overlays of ^{1}H and ^{15}N-HSQC spectra comparing Atox1–Cu–Pt (red) to (**A**) *apo*-Atox1 (black), (**B**) Atox1–Cu (black), and (**C**) Atox1–Pt (black). Peaks showing the largest chemical shift differences are labeled in black and C41 labeled in blue; **D-F:** Combined chemical shift change, relative to *apo*-Atox1 as a function of residue number, caused by (**D**) copper, (**E**) cisplatin, and (**F**) copper followed by cisplatin treatment. The insets show structures of the Atox1 monomer with the Cys residues of the conserved CxxC motif as spheres and chemical shift changes (Δδ > 0.05 ppm) highlighted for Atox1–Cu (**D**, orange), Atox1–Pt (**E**, blue) and Atox1–Cu–Pt (**F**, magenta).

Monte Carlo molecular mechanics results were further refined in density functional theory (DFT) simulations. This yielded two types of stable Cu–S-Pt clusters. The first type contains two copper and one platinum atom (Cu_2PtS_6, Figure 4-3 A) while the second has two copper and two platinum atoms ($Cu_2Pt_2S_8$, Figure 4-3 D). The Cu_2PtS_6 model was qualitatively more stable in the simulations than $Cu_2Pt_2S_8$ and could better resist small perturbations of coordination geometry.

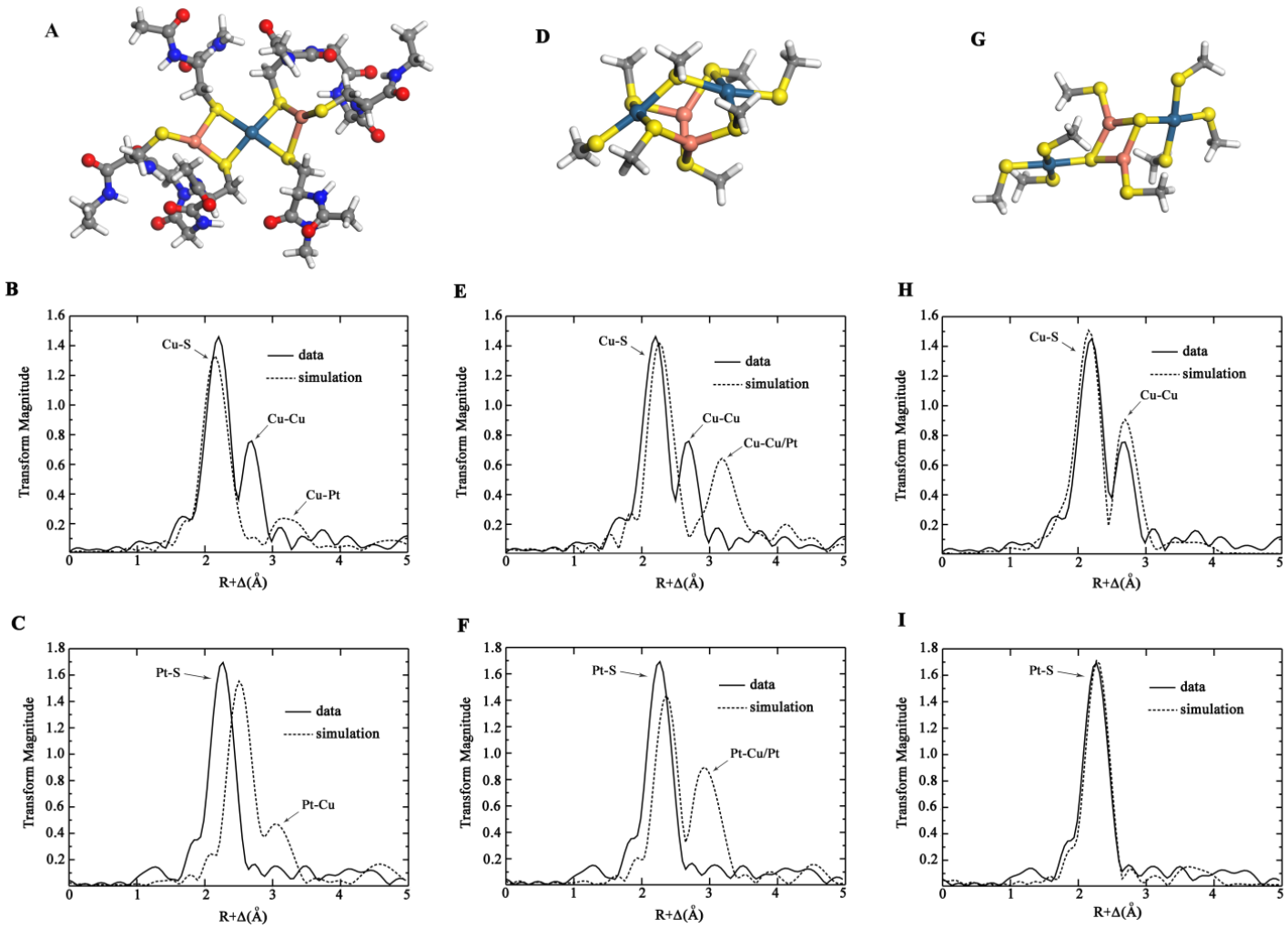


Figure 4-3. Model structures of Atox1–GSH-Cu–Pt complexes and corresponding EXAFS spectra. **A-C:** Cu_2PtS_6 model **D-F:** $Cu_2Pt_2S_8$ model; **G-I:** $Cu_2Pt_2S_{10}$ model; **A, D, G:** Ball-and-stick structure models with copper shown in orange, platinum in navy, sulfur in yellow, and the other atoms in CPK colors. **B, E, H:** Experimental Fourier transformed Cu K-edge EXAFS (solid lines) and corresponding simulated spectra (dashed lines); **C, F, I:** Experimental Fourier transformed Pt L_{III} EXAFS (solid lines) and corresponding simulated spectra (dashed lines).

To further assess the two model structures, simulated EXAFS Fourier transforms were produced (dashed lines in Figure 4-3 B, C, E, F) from the atomic coordinates of the models. Simulated EXAFS did not reproduce the experimental peak corresponding to Cu···Cu interaction at the distance of ~2.7 Å in either case, because both models place copper atoms farther away from each other, at approximately 3.3 Å. The average Cu–S and Pt–S distances in both models were also longer than reported by EXAFS.

To account for the short Cu···Cu distance observed experimentally, we designed a $Cu_2Pt_2S_{10}$ model, which puts copper atoms at 2.7 Å from each other, with two platinum atoms bound through sulfur

bonds to each copper (Figure 4-3 G) Although this model has not been tested using molecular mechanics, it was refined by DFT. Simulated EXAFS for both copper K-edge and platinum L_{III}-edge for $Cu_2Pt_2S_{10}$ model are in good agreement with the experimental data: Fourier transform of copper EXAFS contains ~2.25 Å peak corresponding to Cu···S bond and ~2.7 Å peak corresponding to Cu···Cu interaction (Figure 4-3 H), while Fourier transform of Pt EXAFS accurately reproduces the experimental Pt···S peak.

According to the $Cu_2Pt_2S_{10}$ model, distances between copper and platinum are in the range of 3.7 3.9 Å. At such close distances, the Debye-Waller factor is usually quite high, decreasing the peak amplitude. Moreover, platinum is connected to copper through a single sulfur atom, which allows certain degree of flexibility in the relative positions of copper and platinum, further increasing Debye-Waller factor, and decreasing amplitude of the peak corresponding to Cu···Pt interaction to the noise level (Figure 4-3 H, I). Thus, the $Cu_2Pt_2S_{10}$ model provides the best description of the metal atom arrangement in Atox1 dimer at approximately equimolar ratios of Atox1, copper and platinum.

Under the conditions of chemotherapy, a variety of Atox1 complexes with glutathione, copper and platinum may form in the cell, depending on cisplatin concentration and copper status. A rough estimate based on Atox1 abundance in the cell (***85***), and cisplatin concentration dynamics in blood plasma under the standard chemotherapy regimen (***86***) suggests that the numbers of Atox1 and cisplatin molecules in the cell can be of the same order of magnitude, and therefore structures similar to Cu_2PtS_6 model, which was shown to be stable in simulations, can occur along with $Cu_2Pt_2S_{10}$ and single-metal Atox1 complexes.

Conclusion

Analysis of copper and cisplatin binding to copper chaperone Atox1 under physiological redox conditions revealed the formation of large copper– and platinum–glutathione complexes that are able to transfer both metals to the protein. Binding of copper and cisplatin to Atox1 was found to occur through the formation of copper–sulfur–platinum bridges, where copper is coordinated by three S atoms and platinum by four S atoms. These data offer a new perspective on copper and cisplatin metabolism in the cells where glutathione and Atox1 both participate in platinum transport across the cell, and the metal balance is likely to affect the effectiveness of anticancer chemotherapy with platinum-based drugs.

Chapter 5: Developing and Testing PKA17

Introduction

Assessing protein acidity constant values is important for predicting the structure, stability, and function of proteins. It is therefore beneficial to be able to predict the values of these acidity constants computationally when robust experimental data are not available. Protein pK_a values have to be calculated in aqueous solutions to be biochemically relevant. These acidity constants are proportional to the total free energies of deprotonation. Such a deprotonation-free energy is a sum of the bond breaking energy and the free energy of hydration for the resulting ions. The two components have opposite signs and large magnitudes, often up to hundreds of kcal/mol. The accuracy of calculating the final pK_a value thus depends on reproducing or predicting a very fine balance of energies, since computational predictions need to be accurate within ca. 0.8-1.0 pH units or slightly over 1 kcal/mol in order to be relevant (***87***). This is why robust calculations of protein and other acidity constants remain a very difficult task even in the presence of the computational resources available today. Thus, a number of research groups have applied significant efforts to achieve reliable and accurate results in the computational assessment of protein pK_a values.

Protein pK_a values are proportional to the free energies of deprotonation in an aqueous solution. However, due the difficulty in accurately computing absolute free energies of hydration, it is convenient to calculate pK_a shifts instead of the absolute pK_a values. A pK_a shift is then a measure of the difference in acid dissociation constants between an ionizable group in aqueous solution and the same group in the protein environment. For example, we can use propanoic acid as the reference compound for the aspartic acid residue. Given a reference compound, we want to assess the difference in the deprotonation energy due to the interactions of the residue with the other parts of the protein as opposed to its interactions with bulk solvent.

Historically, the efforts towards creating computational predictions of protein pK_a shifts have been made in several directions. As a starting point, it is natural to apply the Poisson-Boltzmann equation to the calculation of acidity constants (***88***,***89***). Several groups have suggested a number of variations and approximations of this general methodology for pK_a calculations (***90***,***91***,***92***,***93***,***94***).

Other techniques have been applied to the problem as well such as implementing the Poisson-Boltzmann (PB) approach in volume- or surface-based formalisms (***87***). There have been a variety of proposed approaches for optimizing the electrostatic charges used in PB simulations in order to address the need to take into the account the presence of multiple ionizable residues (***91***,***94***). Some research groups have suggested assigning a large dielectric constant (up to 20) to the interior of proteins (***95***). While this

approach led to an overall improvement of the results, it still calculated a number of pK_a values that deviated significantly from experimental values (***87***,***96***). Furthermore, the very physical meaning of such high values of the dielectric constant for the protein interior is not clear.

In order to improve the accuracy of protein acidity constant predictions, some researchers have acknowledged that that conformational changes in the protein in response to the protonation or deprotonation of ionizable residues must be taken into account. The Multi-Conformation Continuum Electrostatic (MCCE) method is particularly interesting: it combines the motion of side-chains with continuum dielectric treatment of solvent and bulk protein (***97***,***98***,***99***). The MCCE method has been used in many applications, including successfully predicting pK_a values for an extensive testing set of several hundred protein residues with AMBER, CHARMM, and PARSE force fields. The prediction results were compared with the experimentally measured acidity constants (***21***).

Some have suggested a number of microscopic techniques with an explicit treatment of solvent (***100***,***101***). Some of these proposed techniques use quantum mechanical representation of the systems While quantum methodology is generally more accurate and rather potent, it unfortunately requires greater amounts of computational resources than non-quantum empirical techniques, and thus its use is currently somewhat limited when protein pK_a calculations are to be carried out (***87***).

Combined quantum mechanical/molecular mechanical (QM/MM) methods can offer a better alternative to pure quantum simulations (***102***,***103***,***104***). Some QM/MM techniques employ complete or partial continuum representation of solvent. One successful example is the applications of constant pH molecular dynamics (CPHMD) simulations (***105***,***106***) In many cases, constant pH simulations approach or match first-principles level of accuracy (***87***). Additionally, CPHMD techniques offer a tool for studying pH-dependent conformational phenomena.

Finally, the scientific community has directed a lot of attention towards the development and application of empirical techniques for evaluating protein pK_a shifts; PROPKA is one of the most successful and widely used examples (***107***). In these techniques, some physical considerations are combined with statistical fitting of descriptors and parameters that predict amino acid pK_a values depending on the environment of the particular ionizable residue. Statistical fitting methods can yield a reasonably high level of accuracy, and such methods implicitly replace any conformational and rotamer sampling that may be needed to account for thermal motion. Such methods are sufficiently accurate in most cases, and they are also very fast and robust when applied to diverse sets of protein residues.

We have created PKA17, a predictor of protein pK_a shifts that has been parametrized on a subset of experimentally known acidity constants of protein residues. One of the distinguishing features of PKA17 is that it employs an extremely coarse-grain model of the protein with each residue represented as only a single particle. This makes the model very simple while also reducing the noise levels since fine variations in atomic positions have no effect on the calculated values of the acidity constants. Another distinguishing feature is the use of a cubic grid model for positioning of the protein residues. Finally, the physical formalism of PKA17 is much simpler than that of PROPKA; the fitting is almost entirely statistical, and we took care to only introduce a minimum number of fittable parameters. At the same time, we believe that this statistical method is defined by the underlying physical principles.

We used 442 protein residues for fitting and benchmarking our model. In spite of its simplicity, we found that PKA17 performed on par with—or slightly better than—PROPKA. We further tested an approach in which we combined the PROPKA and PKA17 results in order to predict protein pK_a shifts. We found that this combined approach is capable of giving better results than either of the techniques alone.

Methods

Mapping of the Protein Geometry to the Grid

Each protein residue is represented by a single particle, the location of which is determined solely by the position of the alpha carbon of the residue. The Cartesian coordinates of the alpha carbons are taken from the input PDB file and then mapped onto nodes of a cubic grid. The side of each grid cell is set to be 5.4 Å. This makes each cubic cell to have a volume that is approximately equal to the average volume of a protein residue.

The process of the geometry mapping by PKA17 is illustrated in Figure 5-1 using chain I from the 1ppf PDB structure. We start with the full PDB structure (a) and then parse it to leave only the alpha carbon locations, each of which represents the whole residue (b). Finally, each of such particles is placed at the nearest node of the cubic lattice with 5.4 Å spacing in each dimension (c). The types of the residues and the connections to the adjacent residues are recorded and retained at this stage. The grid-mapped structure along with the residue type and connectivity information is retained and only this information is used in next stage of the pK_a calculation.

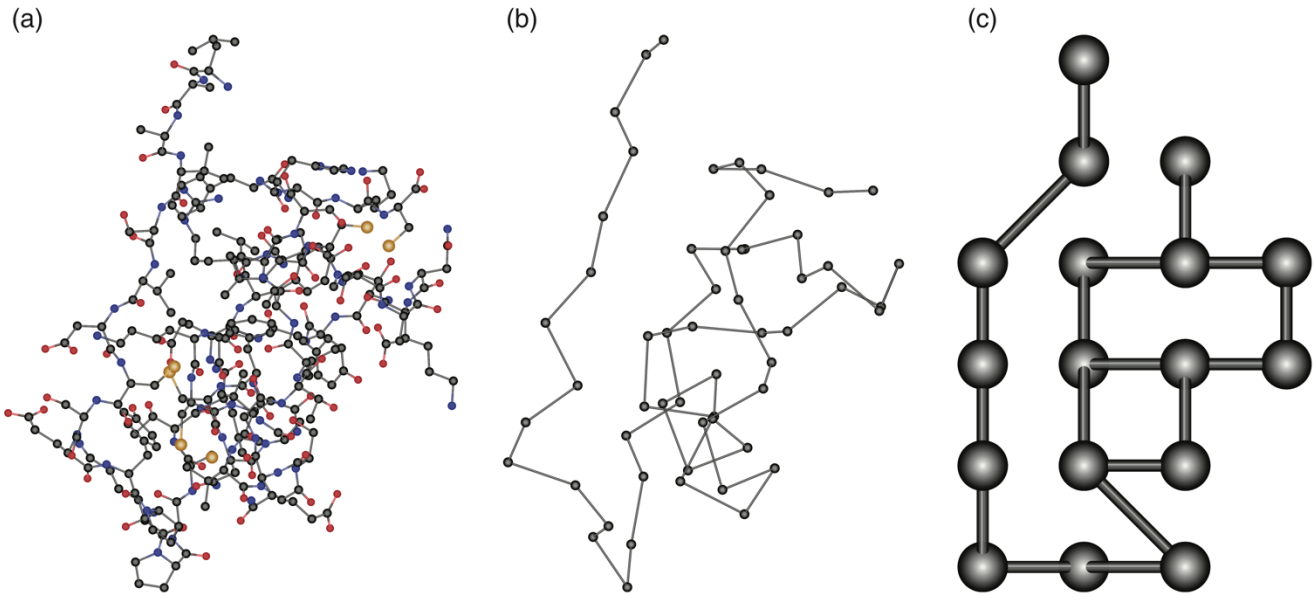


Figure 5-1. Schematic depiction of the process of mapping protein residue coordinates onto the cubic grid. The mapping proceeds from the full atomistic PDB structure (a) to the locations of the alpha carbons (b) and finally to the cubic grid nodes (c).

Selecting Residues and Calculating pK_a Shifts

The current version of PKA17 predicts acidity constants for four types of protein residues: Asp, Glu, His, and Lys. Each of these types has an initial reference pK_a constant A_i^0 where the subscript i denotes the amino acid type. It should be noted that this constant is not intended to correspond to the pK_a value for the residue in any particular protein or peptide, as it is always modified by influence from other residues. The final value of the acidity constant is determined by the following sum:

$$pK_a = A_i^0 + \sum B_{ij} + \sum C_{ij} + \sum D_{ij} + \sum E_{ij} \qquad (22)$$

Figure 5-2 shows graphically how the five coefficients in eq 22 are defined. In detail: Coefficient B_{ij} signifies the shift in the pK_a value of residue A of type i resulting from being directly connected to a residue of type j in the backbone. Coefficient C_{ij} stands for the effect of on the acidity constant of residue A of type i induced by a non-connected residue of type j located just one lattice period l away. Coefficient D_{ij} represents the effect of a diagonally placed residue of type j, and E_{ij} shows the influence of a residue

of type *j* placed at a distance of $\sqrt{3} \cdot l$. Residues that are located farther away do not affect the pK_a value of residue *A*.

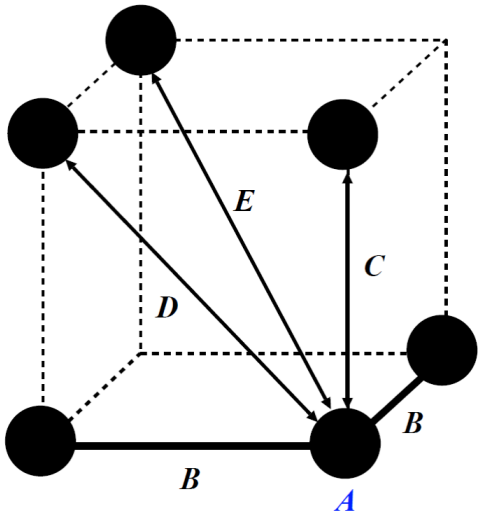


Figure 5-2. Neighboring residues affecting pK_a value of residue *A*.

The values of the coefficients that determine the calculated pK_a value of residue *A* of type *i* depend on the type, *j*, of the influencing residue. Non-connected residues that have been mapped to the same node of the cubic grid as residue *A* are assigned the same pK_a shift coefficients, B_{ij}, as those residues directly connected to *A*.

The above mapping and coefficients are all that determine the values of protein pK_a shifts in the PKA17 framework. Values of all the parameters are found by fitting to experimentally measured values, with further tests on proteins and amino acids that were not included in the fitting set. It should be noted that we are not utilizing any explicit procedure for establishing whether a residue is exposed to the solvent or buried within a protein. However, the shift of the acidity constant does depend on the number of neighboring residues, and thus the effect of exposure to or separation from the solvent is automatically included in an implicit way. Additionally, while the desolvation contribution to the pK_a shift cannot be separated from the other factors contributing to the pK_a shift, it is accounted for as a part of the *B, C, D*, and *E* coefficients in eq 22.

Results and Discussion

In the present version of the PKA17 software, $B_{ij} = C_{ij} = D_{ij} = E_{ij}$ for any *i* and *j*. In other words, any residue of type *j* shifts the pK_a value of residue of type *i* by the same amount if the residues are no farther than $\sqrt{3} \cdot l$ apart. It also does not matter whether the residues are covalently bonded, only the geometric distance between the grid-mapped alpha carbons is used to determine if the residue pairs are neighbors. This was done to avoid overparametrization and related issues with stability and transferability of the results. Naturally, the coefficients are different for different pairs of residue types *i* and *j*. Moreover, it should be emphasized that, in general $B_{ij} \neq B_{ji}$ and the same is true for coefficients *C*, *D*, and *E*.

We used an extensive fitting and testing set of protein residue pK_a values from (***108***,***109***,***110***,***111***,***112***). A complete list of the proteins and residues can be found in the Appendix.
The general fitting procedure was as follows. For each of the ionizable residue types that we considered (Asp, Glu, His, and Lys), we divided the set of experimental pK_a values from the literature into two subsets. The first part was the fitting set. We fitted the parameters A_i^0 and $B_{ij} = C_{ij} = D_{ij} = E_{ij}$ for this residue type to minimize the deviation of the acidity constants calculated with the PKA17 software from the experimental results. The resulting average deviation constituted the first benchmarking result for our fitting.

Then we obtained the leave-one-out (LOO) average unsigned errors. In this procedure, one residue was excluded from the fitting, and the resulting fitted parameters were used to calculate the pK_a value for the excluded residue. The procedure was repeated for all the residues in the set. This allows us to test the sensitivity of our method to small changes in the data set used in parametrization. he next step of our fitting procedure was in applying the parameters derived for the full fitting set to calculate pK_a values of the test set. This was done without any refitting.

Finally, we used the full set (fitting and testing sets together) to fit the final set of the PKA17 parameters and to calculate the LOO average unsigned error. While the LOO result was used as an initial benchmarking measure, the final parameter set that is currently used in the PKA17 software is the one obtained by fitting to the full set of residues. All the calculated errors were compared with those produced with the PROPKA website as PROPKA is one of the most successful and widely used web-based protein pK_a predictors.

Table 5-1. Results of pK_a fitting for Asp, Glu, His, and Lys residues.

	average unsigned error (in pH units) by residue type			
Method	**Asp**	**Glu**	**His**	**Lys**
Fitting set, PROPKA	0.606	0.726	1.038	0.583
Fitting set, PKA17	0.654	0.518	0.722	0.510
Fitting set, PKA17, LOO	0.809	0.632	1.016	0.793
Test set, PROPKA	0.876	0.498	0.872	1.351
Test set, PKA17	0.694	0.479	1.053	1.286
Combined set, PROPKA	0.671	0.669	0.985	0.817
Combined set, PKA17	0.632	0.484	0.730	0.746
Combined set, PKA17, LOO	0.740	0.565	0.914	0.964

Fitting for Asp, Glu, and His Residues

For aspartic acid, we used a fitting set of 105 residues and a testing set of 33 residues. For glutamic acid, we used fitting and testing sets of 101 and 32 residues, respectively. For histidine, a fitting set of 61 residues and a testing set of 28 residues were employed. The results are summarized in Table 5-1 and a complete list of the proteins, residues, and the results can be found in the Appendix.

Fitting PKA17 parameters for Aspartic acid to the fitting set values resulted in an averaged unsigned error of 0.654 pH units, which is a bit higher that the PROPKA 0.606 pH units. The leave-one-out (LOO) average error was 0.809 pH units. While this number is somewhat higher than the PROPKA errors, it should be noted that the PROPKA training set does include some of the residues that were employed in our tests. The performance of this intermediate set of PKA17 parameters for the test subset of the Asp residues was better than that of the PROPKA software, with the average errors being 0.876 and 0.694 pH units for PROPKA and PKA17, respectively. Finally, the fitting to the complete combined set

of Asp residues lead to an average unsigned error of 0.632 pH units as calculated with PKA17. The LOO average error was 0.740 pH units. The PROPKA result is 0.669 pH units. Once again, we need to keep in mind that some of the residues were used in parameter fitting for PROPKA.

For the sets of glutamic acid residues, PKA17 yield lower average unsigned errors of 0.518 pH units for the fitting set and 0.632 pH units for the LOO calculation compared to the average PROPKA error of 0.726 pH units. Evaluating the non-fitted test set produced similar resulting error: 0.479 pH units error for PKA17 and 0.498 units for PROPKA. Using the complete set that includes both the fitting and the testing subsets for the fitting leads to an average PKA17 error of 0.484 pH units and the LOO average unsigned error of 0.565 pH units; both methods resulted in lower average error compared to the corresponding PROPKA average error of 0.671 pH units.

It should be noted explicitly that one of the major reasons for including the leave-one-out errors is the need to test the stability of the resulting PKA17 framework with respect to the fitting data set and its ability to predict acidity constants for residues that are not a part of the fitting set at all. However, it makes sense to minimize the actual final error and going forward, to then use the entire set of residues to produce the finalized version of the PKA17 parameters.

For histidine residues, PKA17 performs better than PROPKA. For the fitting set, applying PKA17 affords an average unsigned error of 0.722 pH units, while the corresponding PROPKA error is 1.038 pH units (~43% higher than PKA11). Even with the leave-one-out (LOO) approach, the PKA17 error is only 1.016 pH units. Interestingly, applying PKA17 to the test set results in a slightly larger error (1.053 pH units) compared PROPKA (1.053 pH units). However, using the complete combined set we PKA17, we obtained average unsigned errors of 0.730 and 0.914 pH units for the fitting and LOO runs, respectively. For comparison, applying PROPKA yields an average unsigned error of 0.985 pH units.

Fitting for Lys Residues

The composition of the fitting and testing sets we compiled for lysine was somewhat different than the sets used for aspartic acid, glutamic acid, and histidine. Most Lys residues for which we could find high-quality experimental data have pK_a values within a relatively narrow range. Our main fitting set is composed of these residues. The testing set (also used as part of the combined set as with Asp, Glu, and His fitting) include engineered mutants that exhibit a much larger range of pK_a values (***113***).This way we covered the broadest range of potential lysine pK_a values that could be encountered. These results are also summarized in Table 5-1.

The initial fitting set contained 57 lysine residues. The average unsigned error in the pK_a values obtained with PROPKA was 0.583 pH units, very similar to the PKA17 average error with this set of 0.510 pH units. The LOO PKA17 result was 0.793 pH units. It should be emphasized again that it is hard to make a direct comparison between PROPKA and PKA17 given that PROPKA parameters were developed with some training/fitting on this particular data set as well.

Calculating lysine pK_a values for the test set of 25 residues yielded average PROPKA and PKA17 errors of 1.351 and 1.286 pH units, respectively. These results are similar, but PKA17 seems to perform slightly better for these structures not employed in the initial fitting. When we use the combined set containing all 82 residues, the average error with PROPKA is 0.817 pH units; the fitting and LOO average with PKA17 are 0.746 and 0.964 pH units, respectively.

PROPKA Fitting Sets for Asp and Glu

The fitting sets which were used to train PROPKA for aspartic and glutamic acid residues are available from the literature (***107***). We have employed these sets in order to produce a more direct comparison of PKA17 and PROPKA. A summary of these can be found in in Table 5-2.

The aspartic acid PROPKA fitting set contains 43 residues (see Appendix for full list). The average unsigned error for the pK_a values for the set computed with PROPKA is 0.503 pH units. When we use the same set for fitting PKA17 parameters, the average error is only 0.299 pH units. The leave-one-out (LOO) procedure resulted in an average error of 0.460 pH—roughly 9% lower than the average PROPKA deviation.

Table 5-2. Results of aspartic and glutamic acid pK_a calculations after fitting to the PROPKA training set (***107***).

	Average unsigned error (in pH units) by residue type	
Method	**Asp**	**Glu**
Fitting set, PROPKA	0.503	0.469
Fitting set, PKA17	0.299	0.331
Fitting set, PKA17, LOO	0.460	0.471
Combined set, PROPKA	0.671	0.669
Combined set, PKA17	0.768	0.603

We then applied the resulting PKA17 parameters to calculating pK_a values of all the 138 aspartic acid residues in our complete combined aspartic acid set. The average PROPKA error (also reported in the above) was 0.671 pH units, while the PKA17 one was slightly higher at 0.768 pH units. Compare this to the average errors produced by PKA17 after fitting to the complete Asp residue set (Table 5-1). The average error for complete fitting was 0.632 pH units, while the LOO procedure lead to an error of 0.740 pH units, close to the 0.768 pH units of unsigned error obtained from fitting to the PROPKA training set.

For glutamic acid, the PROPKA fitting set contained 42 residues. The average unsigned error for these residues that we obtained with PROPKA was 0.469 pH units. Fitting to this set of resides for PKA17 results in an average error of 0.331 pH units, while the LOO protocol gives an error of 0.471 pH units. The resulting parameters were then applied to the full combined set of Glu residues (133 residues in total). For this combined set, the average unsigned PROPKA error in pK_a values is 0.671 pH units. In comparison, the error obtained with PKA17 is 0.603 pH units—about 10% lower.

We believe that these results using the PROPKA fitting set demonstrate the robustness of the PKA17 framework with respect to the fitting protocol and choice of set of fitting set. We've demonstrated that—when both programs are trained and tested on the same data set—the resulting PKA17 parameters allow us to obtain results that are at least as good as those obtained with PROPKA.

Comparative Timing of PROPKA and PKA17

While neither of the programs takes prohibitively long to produce results, we still ran a brief comparison of the required computational time for five representative proteins. The results are shown in Table 5-3. Both programs are rather fast; note that PKA17 is an order of magnitude faster than PROPKA.

Table 5-3. CPU time required for pK_a calculation

PDB ID:	**1ubq**	**2dhc**	**2gga**	**3twy**	**4pyp**
Number of residues	76	310	455	137	504
Number of pK_a values	26	93	108	52	79
Time (s), PROPKA	0.219	1.121	2.947	0.369	1.489
Time (s), PKA17	0.038	0.046	0.062	0.027	0.066

Using PROPKA and PKA17 In Tandem

We have also calculated pK_a values for the complete combined sets, for all four residue types evaluated, by taking the linear combination of the PROPKA and PKA17 results for each residue in equal proportion (0.5 of PROPKA pK_a + 0.5 of PKA17 pK_a). The resulting average unsigned errors are shown in Table 5-4 and Figure 5-3.

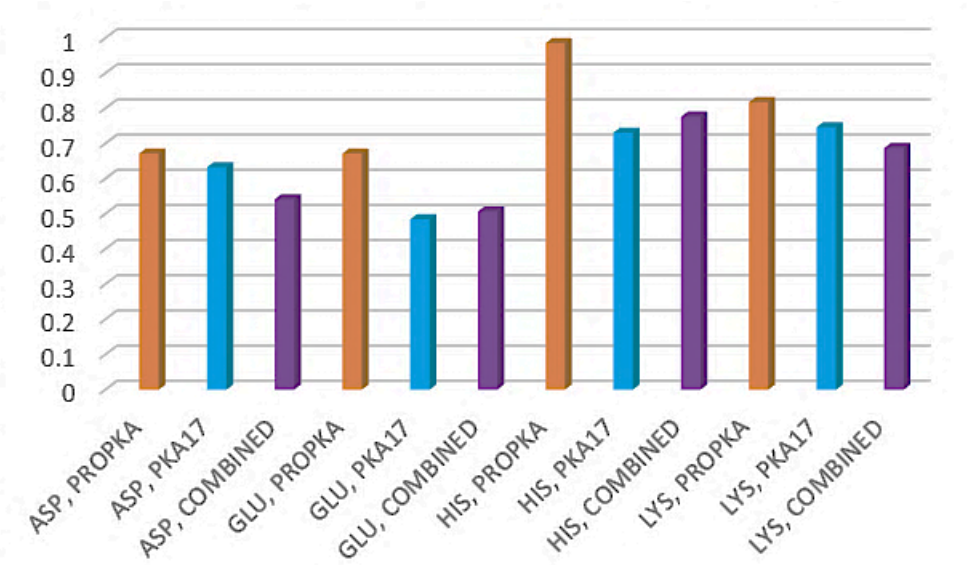


Figure 5-3. Average unsigned errors in pK_a values calculated with PROPKA, PKA17, and as 0.5:0.5 linear combination of the individual PROPKA and PKA17 values (COMBINED).

Table 5-4. Average unsigned errors in pK_a values calculated with PROPKA, PKA17, and as 0.5:0.5 linear combination of the individual PROPKA and PKA17 values.

	average unsigned error (in pH units) by residue type			
	Asp	**Glu**	**His**	**Lys**
PROPKA	0.671	0.671	0.985	0.817
PKA17	0.632	0.484	0.730	0.746
Combined	0.541	0.506	0.776	0.687

There is a clear advantage to using the combined application of PROPKA and PKA17 software. All of the combined PROPKA/PKA17 average errors are lower than those resulting from applying

PROPKA alone. The errors for Asp and Lys residues are also lower than those for applying PKA17 alone. This effect deserves further study as it appears PROPKA and PKA17 tend to err in opposite directions, and their combination provides a more accurate alternative for predicting protein pK_a values.

Benchmarking PKA17 Using an Extensive MCCE Test Set

The Multi-Conformation Continuum Electrostatic (MCCE) method using a Poisson-Boltzmann approach with AMBER, CHARMM, and PARSE force fields has been developed and successfully applied by Alexov and coworkers in order to calculate pK_a values of an extensive set of protein residues (***98***). We have used the same dataset to provide additional validation of our technique by calculating pK_a values with the PKA17 and PROPKA software.

Table 5-5. Average errors of pK_a calculations using the extensive fitting set presented in (***98***)

	average unsigned error (in pH units) by residue type				overall average unsigned error (in pH units)
	Asp	**Glu**	**His**	**Lys**	
MCCE, AMBER [a]	0.473	0.562	0.677	0.479	**0.538**
MCCE, CHARMM [a]	0.464	0.512	0.705	0.504	**0.530**
MCCE, PARCE [a]	0.492	0.519	0.596	0.476	**0.515**
PROPKA	0.600	0.700	0.844	0.542	**0.661**
PKA17	0.572	0.635	0.557	0.507	**0.575**

[a]Data from (***98***)

Table 5-5 show the comparison of accuracy in calculating the acid dissociation constants of residues in the MCCE data set. The data is Table 5-5 are broken down by residue type and the overall accuracy presented in the final column. Note that these results were obtained with no additional fitting of the PKA17 parameters. In almost all cases, the PKA17 formalism performs somewhat worse than the much more sophisticated MCCE technique, but it performs slightly better than the PROPKA suite. The clear exception is histidine calculations in which the PKA17 results are better than those produced by the other techniques. `

Conclusions

We have developed and validated a predictor of protein pK_a values named PKA17. It is based on a coarse-grain grid model of proteins. The pK_a shifts are defined by the residues that are spatially close to the ionizable residue in question. Fitting and validation of our model involved an extensive set of 442 protein residues. Additional benchmarking on a previously proposed extensive set of ionizable protein residues (***98***) confirms the strength of the presented technique. The resulting tool has been deployed with a web-based interface at http://users.wpi.edu/~jpcvitkovic/pka_calc.html.

The results demonstrate that PKA17 performs on par or even somewhat better than the widely used and successful protein pK_a predictor PROPKA. It also requires less computational resources; the computational time needed for PKA17 runs is an order of magnitude lower than the computational time required by PROPKA. Moreover, we have achieved the current level of accuracy with PKA17 while significantly limiting the number of fittable variables in order to avoid any danger of possible overparametrization.

Chapter 6: Conclusions

Parameters for use with the platinum(II) ion, cisplatin, and other Pt(II) coordination compounds and within two force field frameworks—a modified version of fixed-charge OPLS-AA and our polarizable POSSIM force field—were developed and evaluated. The produced seven-site model of Pt(II) and other parameters are intended to be used in protein-metal binding simulations in which—contrary to the common treatment of metal ions in such simulations—the position or even coordination of the ion does not have to be constrained using heuristic routines or operator intervention.

The models have been tested against quantum mechanical and available experimental data for gas-phase complexes, pure liquids, ions solvated in bulk water. The performance has been found adequate, and the second-order polarizable POSSIM technique demonstrated a better ability to reproduce energies of complexes with strong many-body effects present than the fixed-charges modified OPLS-AA. We hope that these new coordination models will permit simulations aimed at obtaining better understanding and prediction of properties of metal ion complexes with proteins and DNA at the all-atom level, even when exact detailed structural information has not been obtained from experimental studies.

We employed our Pt(II) models along with previously developed Cu(I) parameters to investigate, under physiological redox conditions, the binding of platinum to the protein Atox1—a human copper chaperone implicated the resistance mechanism of cisplatin and other platinum antitumor compounds. With our collaborators, we have shown that: (i) Atox1 exists in a dimeric form and binds copper and platinum in the same site formed by the proximal CxxC motifs of the two Atox1 monomers; (ii) copper is coordinated by three and platinum by four sulfur atoms; (iii) glutathione molecules participate in metal binding to Atox1; and (iv) in Atox1–Cu–Pt, copper and platinum do not show a well-defined metal–metal interaction, but are connected through the shared sulfur ligands.

We have developed and benchmarked PKA17, a software package for accurately and quickly calculating protein pK_a values from a supplied protein structure in PDB format. PKA17 employs a highly a coarse-grained lattice model in which each amino acid residue of the input protein structure is represented by only a single particle on a lattice with grid spacing approximately equal to the radius of an "average" amino acid.

To develop PKA17, we carried out parameter fitting using a compilation of 442 aspartic acid, glutamic acid histidine, and lysine residues that had both high-resolution PDB structures and published experimental pK_a values available. The results demonstrate that PKA17 performs on par or even somewhat better than the widely used and successful protein pK_a predictor PROPKA. PKA17 also requires

approximately an order of magnitude less computational resources, as measured by CPU time, than PROPKA Additionally, we have achieved the current level of accuracy with PKA17 while significantly limiting the number of fitting variables in order to avoid any danger of possible overparametrization. We have also shown that the accuracy of PKA17 is reasonably robust with respect to the choice of the fitting set for parametrization, even though some more sophisticated techniques (such and MCCE and quantum mechanics) can yield a higher degree of accuracy of evaluating pK_a shifts of protein residues. Finally, a web interface has been developed for PKA17 and resulting tool has been deployed and is freely accessible (http://users.wpi.edu/~jpcvitkovic/pka_calc.html) and hopefully will prove useful to a wider audience.

References

1. Veluraja, K.; Margulis, C. J. Conformational dynamics of sialyl Lewisx in aqueous solution and its interaction with selectinE. A study by molecular dynamics. *J. Biomol. Struct. Dyn.* **2005,** *23* (1), 101-111.
2. MacDermaid, C. M.; Kaminski, G. A. Electrostatic Polarization Is Crucial for Reproducing pKa Shifts of Carboxylic Residues in Turkey Ovomucoid Third Domain. *J. Phys. Chem. B* **2007,** *111* (30), 9036-9044.
3. Click, T. H.; Kaminski, G. A. Reproducing Basic pKa Values for Turkey Ovomucoid Third Domain Using a Polarizable Force Field. *J. Phys. Chem. B* **2009,** *113* (22), 7844-7850.
4. Kelland, L. The resurgence of platinum-based cancer chemotherapy. *Nat, Rev. Cancer* **2007,** *7* (8), 573-584.
5. Berners-Price, S. J.; Frenkiel, T. A.; Frey, U.; Ranford, J. D.; Sadler, P. J. Hydrolysis products of cisplatin: pKa determinations via[1H, 15N] NMR spectroscopy. *J. Chem. Soc. Chem. Commun.* **1992,** No. 10, 789-791.
6. Siddik, Z. H. Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene* **2003,** *22,* 7265-7279.
7. 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* (6), 887-894.
8. Boal, A. K.; Rosenzweig, A. C. Crystal Structures of Cisplatin Bound to a Human Copper Chaperone. *J. Am. Chem. Soc.* **2009,** *131,* 14196-14197.
9. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001,** *105* (28), 6474-6487.
10. Li, X.; Ponomarev, S. Y.; Sa, Q.; Sigalovsky, D. L.; Kaminski, G. A. *J Comp Chem* **2013,** *34,* 1241-1250.
11. Kaminski, G. A.; Ponomarev, S. Y.; Liu, A. B. Polarizable simulations with second order interaction model--force field and software for fast polarizable calculations: parameters for small model systems and free energy calculations. *J. Chem. Theory Comput.* **2009,** *5* (11), 2935-2943.

12. Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. *J Comp Chem* **2002,** *23,* 1515-1531.

13. Schrödinger, LLC. *Jaguar, v7.6;* New York, NY, 2009.

14. Kaminski, G. A.; Maple, J. R.; Murphy, R. B.; Braden, D. A.; Friesner, R. A. *J Chem Theory Comput* **2005,** *1,* 248-254.

15. Li, X.; Ponomarev, S. Y.; Sigalovsky, D. L.; Cvitkovic, J. P.; Kaminski, G. A. POSSIM: Parameterizing Complete Second-Order Polarizable Force Field for Proteins. *J. Chem. Theory Comput.* **2014,** *10* (11), 4896-4910.

16. Dmitriev, O. Y. Mechanism of tumor resistance to cisplatin mediated by the copper transporter ATP7B. *Biochem. Cell Biol.* **2011,** *89* (2), 138-147.

17. Jamieson, E. R.; Lippard, S. J. Structure, recognition, and processing of cisplatin-DNA adducts. *Chem. Rev.* **1999,** *99,* 2467– 2498.

18. Bischin, C.; Lupan, A.; Taciuc, V.; Silaghi-Dumitrescu, R. Interactions between proteins and platinum-containing anti-cancer drugs. *Mini-Reviews in Medicinal Chemistry* **2011,** *11,* 214-224.

19. Dolgova, N.; Nokhrin, S.; Yu, C.; George, G.; Dmitriev, O. Copper chaperone Atox1 interacts with the metal-binding domain of Wilson's disease protein in cisplatin detoxification. *Biochemical Journal* **2013,** *454,* 147-156.

20. Yao, S.; Plastaras, J. P.; Marzilli, L. G. A Molecular Mechanics AMBER-Type Force Field for Modeling Platinum Complexes of Guanine Derivatives. *Inorg. Chem.* **1994,** *33* (26), 6061-6077.

21. Wang, X.; Li, C.; Wang, Y.; Chen, G. Interaction of classical platinum agents with the monomeric and dimeric Atox1 proteins: a molecular dynamics simulation study. *Int. J. Mol. Sci.* **2013,** *15* (1), 75-99.

22. Banci, L. Molecular dynamics simulations of metalloproteins. *Curr. Opinion Chem. Biology* **2003,** *7,* 143-149.

23. Xu, L.; Wang, X.; Shan, S.; Wang, X. Characterization of the Polymorphic Sates of Copper(II)-Bound Ab(1-16) Peptides by Computational Simulations. *J. Comput. Chem.* **2013,** *34,* 2524-2536.

24. Huang, J.; Devereux, M.; Hofmann, F.; Meuwly, M. Computational Organometallic Chemistry with Force Fields. In *Computational Organometallic Chemistry, ;* Wiest, O., Wu, Y., Eds.; Springer, 2012; pp 19-46.

25. Deeth, R. J. General Molecular Mechanics Method for Transition Metal Carboxylates and its Application to Multiple Coordination Modes in Mono- and Dinuclear Mn(II) Complexes. *Inorg. Chem.* **2008,** *47* (15), 6711-6725.

26. Piquemal, J. P.; Gresh, N.; Giessner-Prettre, C. Improved Formulas for the Calculation of the Electrostatic Contribution to the Intermolecular Interaction Energy from Multipolar Expansion of the Electronic Distribution. *J. Phys. Chem. A* **2003,** *107* (48), 10353-10359.

27. Piquemal, J. P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darde, T. A. Toward a force field based on density fitting. **2006,** *124,* 104101.

28. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. . J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.03;* Gaussian, Inc.: Wallingford, CT, 2009.

29. Paschoal, D.; Marcial, B. L.; Lopes, J. F.; De Almeida, W. B.; Dos Santos, H. F. *J Comp Chem* **2012,** *33,* 2292-2302.

30. Ponomarev, S. Y.; Kaminski, G. A. Polarizable simulations with second-order interaction model (POSSIM) force field: developing parameters for alanine pep Simulations with Second-Order Interaction Model (POSSIM) Force Field: Developing Parameters for Alanine Peptides and Protein Backbone. *J. Chem. Theory Comput.* **2011,** *7* (5), 1415–1427.

31. Ponomarev, S. Y.; Click, T. H.; Kaminski, G. A. Electrostatic Polarization Is Crucial in Reproducing Cu(I) Interaction Energies and Hydration. *J. Phys. Chem. B* **2011,** *115* (33), 10079–10085.

32. Rizzo, R. C.; Jorgensen, W. L. *J Am Chem Soc* **1999,** *121,* 4827-4836.

33. Haar, L.; Gallagher, J. S. *J Phys Chem Ref Data* **1978,** *7,* 635.

34. Jensen, K. P.; Jorgensen, W. L. *J Chem Theory Comput* **2006,** *2,* 1499-1509.

35. Chandrasekhar, J.; Spellmeyer, D. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984,** *106* (4), 903-910.

36. Masia, M. Estimating Chloride Polarizability in a Water Solution. *J Phys. Chem. A* **2013,** *117* (15), 3221-3226.

37. Jungwrith, P.; Tobias, D. J. Chloride Anion on Aqueous Clusters, at the Air−Water Interface, and in Liquid Water: Solvent Effects on Cl- Polarizability. *J. Phys. Chem. A* **2002,** *106* (2), 379-383.

38. Hättig, C.; Heß, B. A. TDMP2 calculation of dynamic multipole polarizabilities and dispersion coefficients for the halogen anions F−, Cl−, Br− and I−. *J. Chem. Phys.* **1998,** *108* (10), 3863-3870.

39. 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* (2), 926-935.

40. Likholyot, A.; Hovey, J. K.; Seward, T. M. Experimental and theoretical study of hydration of halide ions. *Geochim. Cosmochim. Acta* **2005,** *69* (12), 2949-2958.

41. Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. *J. Phys. Chem. A* **1998,** *102* (40), 7787-7794.

42. Sieck, L. W.; Meot-Ner, M. Ionic Hydrogen Bond and Ion Solvation. 8. RS–···HOR Bond Strengths: Correlation with Acidities. *J. Phys. Chem.* **1989,** *93* (4), 1586–1588.

43. Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous Solvation Free Energies of Ions and Ion−Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. *J. Phys. Chem. B* **2006,** *110* (32), 16066-16081.

44. Wagman, D. D. The NBS Tables of Chemical Thermodynamic Properties. *J. Phys. Chem. Ref. Data* **1982,** *11* (2), 1-151.

45. Hofer, T. S.; Randolf, B. R.; Rode, B. M.; Persson, I. The hydrated platinum(ii) ion in aqueous solution—a combined theoretical and EXAFS spectroscopic study. *Dalton Trans.* **2009,** No. 9, 1512-1515.

46. Horn, D.; Barrientos, A. Mitochondrial Copper Metabolism and Delivery to Cytochrome c Oxidase. *IUBMB Life* **2008,** *60* (7), 421-429.

47. Parvez, S.; Kang, M.; Chung, H. S.; Bae, H. Naturally Occurring Tyrosinase Inhibitors: Mechanism and Applications in Skin Health, Cosmetics and Agriculture Industries. *Phytotherapy Research* **2007,** *21,* 805-816.

48. Opazo, C. M.; Greenough, M. A.; Bush, A. I. Copper: from neurotransmission to neuroproteostasis. *Frontiers in aging neuroscience* **2014,** *6,* 143.

49. Watanabe, K.; Shibuya, S.; Ozawa, Y.; Nojiri, H.; Izuo, N.; Yokote, K.; Shimizu, T. Superoxide Dismutase 1 Loss Disturbs Intracellular Redox Signaling, Resulting in Global Age-Related Pathological Changes. *BioMed Research International* **2014,** *2014*.

50. Lutsenko, S.; Barnes, N. L.; Bartee, M. Y.; Dmitriev, O. Y. Function and regulation of human copper-transporting ATPases. *Physiol. Rev.* **2007,** *87* (3), 1011-1046.

51. Ahuja, A.; Dev, K.; Tanwar, R. S.; Selwal, K. K.; Tyagi, P. K. Copper mediated neurological disorder: Visions into amyotrophic lateral sclerosis, Alzheimer and Menkes Disease. *J. Trace Elem. Med. Biol.* **2015,** *29,* 11-23.

52. Myhre, O.; Utkilen, H.; Duale, N.; Brunborg, G.; Hofer, T. Metal Dyshomeostasis and Inflammation in Alzheimer's and Parkinson's Diseases: Possible Impact of Environmental Exposures. *Oxid. Med. Cell. Longevity* **2013,** *2013*.

53. Brewer, G. J. The promise of copper lowering therapy with tetrathiomolybdate in the cure of cancer and in the treatment of inflammatory disease. *J. Trace Elem. Med. Biol.* **2014,** *28* (4), 372– 378.

54. Rae, T. D.; Schmidt, P. J.; Pufahl, R. A.; Culotta, V. C.; O'Halloran, T. V. Undetectable Intracellular Free Copper: the Requirement of a Copper Chaperone for Superoxide Dismutase. *Science* **1999,** *284,* 805-808.

55. 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* (22), 14298-14302.

56. Apps, M. G.; Choi, E. H. Y.; Wheate, N. J. The state-of-play and future of platinum drugs. *Endocr.-Relat. Cancer* **2015,** *22,* R219– R233.

57. Dolgova, N. V.; Olson, D.; Lutsenko, S.; Dmitriev, O. The soluble metal-binding domain of the copper transporter ATP7B binds and detoxifies cisplatin. *Biochem. J.* **2009,** *419,* 51– 56.

58. Ma, G.; Wu, Q.; Wu, X.; Arnesano, F.; Natile, G.; Sletten, E.; Liu, Y. The reaction of a platinated methionine motif of CTR1 with cysteine and histidine is dependent upon the type of precursor platinum complex. *J. Inorg. Biochem.* **2005,** *153,* 239–246.

59. Nguyen, T. H.; Arnesano, F.; Scintilla, S.; Rossetti, G.; Ippoliti, E.; Carloni, P.; and 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* (8), 2912–2920.

60. Galliani, A.; Losacco, M.; Lasorsa, A.; Natile, G.; Arnesano, F. Cisplatin handover between copper transporters: the effect of reducing agents. *J Biol Inorg Chem* **2014**.

61. Safaei, R.; Otani, S.; Larson, B. J.; Rasmussen, M. L.; Howell, S. B. Transport of cisplatin by the copper efflux transporter ATP7B. *Molecular Pharmacology* **2008,** *73* (2), 461-468.

62. Hua, H.; Günther, V.; Georgiev, O.; Schaffner, W. Distorted copper homeostasis with decreased sensitivity to cisplatin upon chaperone Atox1 deletion in Drosophila. *Biometals* **2011,** *24* (3), 445–453.

63. 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* (3), 333-341.

64. Beaino, W.; Guo, Y.; Chang, A. J.; Anderson, C. J. Roles of Atox1 and p53 in the trafficking of copper-64 to tumor cell nuclei: implications for cancer therapy. *J. Biol. Inorg. Chem.* **2014,** *19* (3), 423-438.

65. 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.

66. Palm-Espling, M. E.; Andersson, C. D.; Bjorn, E.; Linusson, A.; Wittung-Stafshede, P. Determinants for Simultaneous Binding of Copper and Platinum to Human Chaperone Atox1: Hitchhiking not Hijacking. *PloS One* **2013,** *8* (7), e70473.

67. Xi, Z.; Guo, W.; Tian, C.; Wang, F.; Liu, Y. Copper binding promotes the interaction of cisplatin with human copper chaperone Atox1. *Chem. Commun.* **2013,** *49,* 11197-11199.

68. Deponte, M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochim. Biophys. Acta, Gen. Subj.* **2013,** *1830* (5), 3217–3266.

69. Dolgova, N. V.; Yu, C.; Cvitkovic, J. P.; Hodak, M.; Nienaber, K. H.; Summers, K. L.; Cotelesage, J. J.; Bernholc, J.; Kaminski, G. A.; Pickering, I. J.; George, G. N. Binding of Copper and Cisplatin to Atox1 Is Mediated by Glutathione through the Formation of Metal–Sulfur Clusters. *Biochemistry* **2017,** *56* (24), 3129–3141.

70. Wernimont, A. K.; Huffman, D. L.; Lamb, A. L.; O'Halloran, T. V.; Rosenzweig, A. C. Structural basis for copper transfer by the metallochaperones for the Menkes/Wilson disease proteins. *Nature Structural Biology* **2000,** *7,* 766-771.

71. 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.

72. Jorgensen, W. L.; Tirado-Rives, J. Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. *J. Comput. Chem.* **2005,** *26,* 1689–1700.

73. Click, T. H.; Ponomarev, S. Y.; Kaminski, G. A. Importance of electrostatic polarizability in calculating cysteine acidity constants and copper(I) binding energy of Bacillus subtilis CopZ. *J. Comput. Chem.* **2012,** *33,* 1142–1151.

74. Cvitkovic, J. P.; Kaminski, G. A. Developing multisite empirical force field models for Pt(II) and cisplatin. *J. Comput. Chem.* **2017,** *38,* 161–168.

75. Hodak, M.; Lu, W.; Bernholc, J. Hybrid ab initio Kohn-Sham density functional theory/frozen-density orbital-free density functional theory simulation method suitable for biological systems. *J. Chem. Phys.* **2008,** *129* (1), 014101–014109.

76. Bernholc, J.; Hodak, M.; Lu, W. Recent developments and applications of the real-space multigrid method. *J. Phys.: Condens. Matter* **2008,** *20,* 294205.

77. Briggs, E. L.; Sullivan, D. J.; Bernholc, J. Real-space multigrid-based approach to large-scale electronic structure calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996,** *54* (20), 14362–14375.

78. Hodak, M.; Wang, S.; Lu, W.; Bernholc, J. Implementation of ultrasoft pseudopotentials in large-scale grid-based electronic structure calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007,** *76,* 085108.

79. Schafer, F. Q.; Buettner, G. R. Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radical Biol. Med.* **2011,** *30* (11), 1191–1212.

80. Pickering, I. J.; George, G. N.; Dameron, C. T.; Kurz, B.; Winge, D. R.; Dance, I. G. X-ray Absorption Spectroscopyof Cuprous-Thiolate Clusters in Proteins and Model Systems. *J. Am. Chem. Soc.* **1993,** *115,* 9498-9505.

81. Zhang, L.; Pickering, I. J.; Winge, D. R.; George, G. N. X-ray absorption spectroscopy of cuprous-thiolate clusters in Saccharomyces cerevisiae metallothionein. *Chem. Biodiversity* **2008,** *5,* 2042–2049.

82. Ralle, M.; Lutsenko, S.; Blackburn, N. X-ray absorption spectroscopy of the copper chaperone HAH1 reveals a linear two-coordinate cu(i) center capable of adduct formation with exogenous thiols and phosphines. *J. Biol. Chem.* **2003,** *278* (25), 23163–23170.

83. Hussain, F.; Rodriguez-Granillo, A.; Wittung-Stafshede, P. Lysine-60 in Copper Chaperone Atox1 Plays an Essential Role in Adduct Formation with a Target Wilson Disease Domain. *J. Am. Chem. Soc.* **2009,** *131,* 16371-16373.

84. Xiao, Z.; Brose, J.; Schimo, S.; Ackland, S. M.; La Fontaine, S.; Wedd, A. G. Unification of the copper(I) binding affinities of the metallo-chaperones Atx1, Atox1, and related proteins: detection probes and affinity standards. *J. Biol. Chem.* **2011,** *286* (13), 11047–11055.

85. Beck, M.; Schmidt, A.; Malmstroem, J.; Claassen, M.; Ori, A.; Szymborska, A.; Herzog, F.; Rinner, O.; Ellenberg, J.; Aebersold, R. The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **2011,** *7,* 549.

86. Himmelstein, K. J.; Patton, T. F.; Belt, R. J.; Taylor, S.; Repta, A. J.; Sternson, L. A. Clinical kinetics on intact cisplatin and some related species. *Clin. Pharmacol. Ther.* **1981,** *29* (5), 658-644.

87. Alexov, E.; Mehler, E. L.; Baker, N.; Baptista, A. M.; Huang, Y.; Milletti, F.; Nielsen, J. E.; Farrell, D.; Carstensen, T.; Olsson, M. H. M.; Shen, J. K.; Warwicker, J.; Williams, S.; Word, J. M. Progress in the prediction of pKa values in proteins. *Proteins* **2011,** *79* (12), 3260−3275.

88. Tanford, C.; Kirkwood, J. G. Theory of Protein Titration Curves. I. General Equations for Impenetrable Spheres. *J. Am. Chem. Soc.* **1957,** *79* (20), 5333–5339.

89. Bashford, D.; Karplus, M. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry* **1990,** *29* (44), 10219–10225.

90. Lim, C.; Bashford, D.; Karplus, M. Absolute pKa calculations with continuum dielectric methods. *J. Phys. Chem. 1991 95(14)* **1991,** *95* (14), 5610–5620.

91. Gilson, M. K. Multiple-site titration and molecular modeling: Two rapid methods for computing energies and forces for ionizable groups in proteins. *Proteins* **1993,** *15* (3), 266–282.

92. Potter, M.; Gilson, M.; McCammon, J. Small molecule pKa prediction with continuum electrostatics calculations. *J. Am. Chem. Soc.* **1994,** *116* (20), 10298–10299.

93. Antosiewicz, J.; McCammon, J.; Gilson, M. Prediction of pH-dependent Properties of Proteins. *J. Mol. Biol.* **1994,** *238* (3), 415–436.

94. Karshikoff, A. A simple algorithm for the calculation of multiple site titration curves. *Protein Eng.* **1995,** *8* (3), 243–248.

95. Antosiewicz, J.; Briggs, J.; Elcock, A.; Gilson, M.; McCammon, J. Computing ionization states of proteins with a detailed charge model. *J. Comp. Chem.* **1996,** *17* (14), 1633–1644.

96. Teixeira, V. H.; Cunha, C. A.; Machuqueiro, M.; Oliveira, A. S.; Victor, B. L.; Soares, C. M.; Baptista, A. M. On the Use of Different Dielectric Constants for Computing Individual and Pairwise Terms in Poisson−Boltzmann Studies of Protein Ionization Equilibrium. *J. Phys. Chem. B 2005 109(30)* **2005,** *109* (30), 14691–14706.

97. Georgescu, R.; Alexov, E.; Gunner, M. Combining conformational flexibility and continuum electrostatics for calculating pKas in proteins. *Biophys. J.* **2002,** *83* (4), 1731–1748.

98. Wang, L.; Li, L.; Alexov, E. pKa predictions for proteins, RNAs, and DNAs with the Gaussian dielectric function using DelPhi pKa. **2015,** *83* (12), 2186–2197.

99. Song, Y.; Mao, J.; Gunner, M. R. MCCE2: Improving protein pKa calculations with extensive side chain rotamer sampling. *Comput. Chem.* **2009,** *30* (14), 2231–2247.

100. Mehler, E. L. The Lorentz-Debye-Sack theory and dielectric screening of electrostatic effects in proteins and nucleic acids. *Theor. Comput. Chem.* **1996,** *3,* 371–405.

101. Schulz, C.; Warshel, A. What are the dielectric "constants" of proteins and how to validate electrostatic models?. *Proteins* **2011,** *44* (4), 400–417.

102. Li, H.; Robertson, A. D.; Jensen, J. H. The determinants of carboxyl pKa values in turkey ovomucoid third domain. *Proteins* **2004,** *55* (3), 689–704.

103. Jensen, J. H.; Li, H.; Robertson, A. D.; Molina, P. A. Prediction and Rationalization of Protein pKa Values Using QM and QM/MM Methods. *J. Phys. Chem. A* **2005,** *109* (3), 6634–6643.

104. Schaefer, P.; Riccardi, D.; Cui, Q. Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules. *J. Chem. Phys.* **2005,** *123* (1), 014905.

105. Lee, M. S.; Salsbury, F. R. B. C. L. Constant-pH molecular dynamics using continuous titration coordinates. *Proteins* **2004,** *54* (4), 738–752.

106. Wallace, J. A.; Shen, J. K. Continuous constant pH molecular dynamics in explicit solvent with pH-based replica exchange. *J. Chem. Theory. Comput.* **2011 ,** *7* (8), 2617–2629.

107. Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory. Comput.* **2010,** *7* (2), 525-537.

108. Castaneda, C. A.; Fitch, C. A.; Majumdar, A.; Khangulov, V.; Schlessman, J. L.; Garcia-Moreno, B. E. Molecular determinants of the pKa values of Asp and Glu residues in staphylococcal nuclease. *Proteins* **2009,** *77* (3), 570-588.

109. Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci.* **2009,** *18* (1), 247-251.

110. Arthur, E. J.; Yesselman, J. D.; Brooks, C. L. Predicting extreme pKa shifts in staphylococcal nuclease mutants with constant pH molecular dynamics. *Proteins* **2011,** *79* (12), 3276-3286.

111. Hiebler, K.; Lengyel, Z.; Castaneda, C. A.; Makhlynets, O. V. Functional tuning of the catalytic residue pKa in a de novo designed esterase. *Proteins* **2017,** *85* (9), 1656-1665.

112. Wu, X.; Lee, J.; Brooks, B. R. Origin of pKa Shifts of Internal Lysine Residues in SNase Studied Via Equal-Molar VMMS Simulations in Explicit Water. *J. Phys. Chem. B* **2017,** *121* (15), 3318-3330.

113. Wu, X.; Lee, J.; Brooks, B. R. Origin of pKa Shifts of Internal Lysine Residues in SNase Studied Via Equal-Molar VMMS Simulations in Explicit Water. *J. Phys. Chem. B* **2017,** *121* (15), 3318-3330.

114. Appleton, T. G.; Connor, J. W.; Hall, J. R.; Prenzler, P. D. NMR Study of the Reactions of the cis-Diamminediaquaplatinum(II) Cation with Glutathione and Amino Acids Containing a Thiol Group. *Inorganic Chemistry* **1989,** *28,* 2030-2037.

115. Badarau, A.; Dennison, C. Copper Trafficking Mechanism of CXXC-Containing Domains: Insight from the pH-Dependence of Their Cu(I) Affinities. *J. Am. Chem. Soc.* **2011,** *133,* 2983-2988.

116. Xi, Z.; Shi, C.; Tian, C.; Liu, Y. Conserved residue modulates copper-binding properties through structural dynamics in human copper chaperone Atox1. *Metallomics* **2013,** *5,* 1566-1573.

117. Brose, J.; La Fontaine, S.; Wedd, A. G.; Xiao, Z. Redox sulfur chemistry of the copper chaperone Atox1 is regulated by the enzyme glutaredoxin 1, the reduction potential of the glutathione couple GSSG/2GSH and the availability of Cu(I). *Metallomics* **2014**.

118. Morgan, C. T.; Tsivkovskii, R.; Kosinsky, Y. A.; Efremov, R. G.; Lutsenko, S. The distinct functional properties of the nucleotide-binding domain of ATP7B, the human copper-transporting ATPase: analysis of the Wilson disease mutations E1064A, H1069Q, R1151H, and C1104F. *J. Biol. Chem.* **2004,** *279,* 36363-36371.

119. Ciriolo, M. R.; Desideri, A.; Paci, M.; Rotilio, G. Reconstitution of Cu,Zn-superoxide dismutase by the Cu(I)-glutathione complex. *J. Biol. Chem.* **1990,** *265,* 11030-11034.

120. Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Kozyreva, T.; Zovo, K.; Palumaa, P. Affinity gradients drive copper to cellular destinations. *Nature* **2010,** *465,* 645-650.

121. Miras, R.; Morin, I.; Jacquin, O.; Cuillel, M.; Guillain, F.; Mintz, E. Interplay between glutathione, Atx1 and copper. 1. Copper(I) glutathionate induced dimerization of Atx1. *JBIC J. Biol. Inorg. Chem.* **2008,** *13* (2), 195-205.

122. Palm, M. E.; Weise, C. F.; Lundin, C.; Wingsle, G.; Nygren, Y.; Björn, E.; Naredi, P.; Wolf-Watz, a. W.-S. P. Cisplatin binds human copper chaperone Atox1 and promotes unfolding in vitro. *PNAS* **2011,** *108* (17), 6951-6956.

123. Poger, D.; Fillaux, C.; Miras, R.; Crouzy, S.; Delangle, P.; Mintz, E.; Den Auwer, C.; Ferrand, M. Interplay between glutathione, ATX1 and copper: X-ray absorption spectroscopy determination of Cu(I) environment in an Atx1 dimer. *JBIC, J. Biol. Inorg. Chem.* **2008,** *13,* 1239-1248.

124. Pushie, M. J.; Zhang, L.; Pickering, I. J.; George, G. N. The fictile coordination chemistry of cuprous-thiolate sites in copper chaperones. *BIOchim. Biophys. Acta, Bioenerg.* **2012,** *1817,* 938-947.

125. Rubino, J. T.; Franz, K. J. Coordination chemistry of copper protens: how nature handles toxic cargo for essential function. *J. Inorg. Biochem.* **2012,** *107,* 129-143.

126. Tanchou, V.; Gas, F.; Urvoas, A.; Cougouluegne, F.; Ruat, S.; Averseng, O.; Quemeneur, E. Copper-mediated homo-dimerisation for the HAH1 metallochaperone. *Biochem. Biophys. Res. Commun.* **2004,** *325,* 388-394.

127. Narindrasorasak, S.; Zhang, X.; Roberts, E. A.; Sarkar, B. Comparative analysis of metal binding characteristics of copper chaperone proteins, Atx1 and Atox1. *Bioinorg. Chem. Appl.* **2004,** *2* (1-2), 105-123.

128. Lim, C.; Bashford, D.; Karplus, M. Absolute pKa calculations with continuum dielectric methods. *J. Phys. Chem.* **1991,** *95* (14), 5610–5620.

129. Corazza, A.; Harvey, I.; Sadler, P. J. 1H,13C-NMR and X-ray absorption studies of copper(I) glutathione complexes. *Eur. J. Biochem.* **996,** *236,* 697-705.

Appendices

Table A-1. Platinum (II) and cisplatin parameter values used in this work.

Atom	charge, e^-	Lennard-Jones ε, kcal/mol	Lennard-Jones σ, Å	Inverse polarizability α^{-1}, Å^{-3}	Pt–X distance, Å
Modified OPLS					
Pt in 7 site Pt^{2+}	0.00	3.14	0.015		
X_{ax} in 7 site Pt^{2+}	0.00	0.00	0.000		1.3
X_{eq} in 7 site Pt^{2+}	0.50	0.00	0.000		1.3
S in CH_3S^-	–0.900	4.50	0.066		
C in CH_3S^-	–0.280	4.20	0.300		
H in CH_3S^-	0.060	2.50	0.030		
Cl^-	–1.00	4.417	0.118		
N in NH_3	–1.02	3.42	0.170		
H in NH_3	0.34	0.00	0.000		
POSSIM					
Pt in 7 site Pt^{2+}	0.00	2.99	0.015	–	
X_{ax} in 7 site Pt^{2+}	0.09	0.00	0.000	–	0.87
X_{eq} in 7 site Pt^{2+}	0.455	0.00	0.000	–	0.87
Cl^-	–1.00	4.81	0.080	0.20	
N in NH_3	–0.87	3.465	0.120	0.622	
H in NH_3	0.29	0.00	0.000	–	
S in CH_3S^-	–0.925	4.51	0.200	0.39	
C in CH_3S^-	–0.255	3.50	0.066	0.5069	
H in CH_3S^-	0.060	2.50	0.030	–	