Modeling Charge Transport Through Biological Molecules with Transition Metals

William Livernois

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Anant M. P. Anantram, Chair

Ting Cao

Arindam Kumar Das

Program Authorized to Offer Degree

Electrical and Computer Engineering

University of Washington

Abstract

Modeling Charge Transport Through Biological Molecules with Transition Metals

William Livernois

Chair of the Supervisory Committee:

Anant M. P. Anantram

Electrical and Computer Engineering

In this work the electronic structure of nucleic acids (modified DNA) and proteins are modeled with a focus on incorporation of transition metals. Building on previous work, experimental data has been combined with molecular mechanics methods, density functional theory calculations, and Green's function methods to model charge transport through these structures as well as electronic behavior in varying environments. Addition of transition metals poses new challenges at each one of these steps, owing to the d-orbital interactions with coordinating ligands and the possibility of ground states with unpaired spin, which was explored in the analysis of two different systems. The

first system was a metal-modified DNA structure, using modified base pairs to intercalate silver, mercury, and gold ions. Electronic structure calculations show that the intercalated metals create isolated valence states that can assist electron transport along the DNA strand. For structure generation in this system, a combination of ab-initio methods and generated molecular dynamics force fields were used to model interactions with the coordinated metal. The second system modeled was the OmcS cytochrome, which contains heme-coordinated Fe centers that can exist at multiple oxidation states. Therefore, a spin-dependent model must be used, resulting in spin-dependence of charge transport in the oxidized state, shown in both hopping models and Green's function decoherent models. A non-collinear spin model was also developed, showing that the spin-orbit coupling from the contact influences the spin-dependence of transport.

Acknowledgements

First, I want to thank my advisor, Professor Anantram, who dealt with my antics for five years and supported me even when I was a Troublemaker, as well as all the ECE staff, who made sure I never fell through the cracks even while facing a global pandemic and huge cuts to scientific research funding in the US. I am so grateful to my loving partner Alex Bernson, who supported me indirectly outside of academia and directly with professional and detailed edits of every single publication (including this one) since we met. Thanks to both of my parents for encouraging my journey as a scientist, buying me textbooks that I picked out at garage sales and cleaning up my early science experiments done with caustic household chemicals.

I have had so many wonderful collaborators over the years, both locally and abroad, that have made all this work possible. In our lab group, I am grateful to the mentorship from Prof. Hashem Mohammad and Dr. Busra Demir, as well as the fantastic collaborations I have had with Arpan De, Simon Cao, Olaiyan Alolaiyan, and Sachal Shaik. Many thanks to our experimental collaborations with Dr. Simon Vecchioni and the DNA Nanotechnology laboratory research group at NYU and Professor Nikhil Malvankar and his research group at Yale University.

I am also so grateful to all the local grassroots climate justice organizations that kept me sane as we face our unjust present and uncertain future. I want to especially call out Extinction Rebellion and Scientist Rebellion, to whom I gave much of my heart and soul over the years, as well as our local union UAW 4121, who supported me and my peers as we struggled with life as academic student employees. It was in these groups that I found hope for our collective future.

I would like to thank the friendly staff at Saigon Deli for serving the best Vietnamese food in the area at a reasonable price point, their Bahn Mi sandwiches were critical to my survival as a PhD student at UW. I would also like to thank Beyonce, Doechii, Billie Eilish, Japanese Breakfast, AJJ, and Thot Squad for creating music that kept me sane during these five years. Thanks to June Isenhart, who coined the term “trashteria” on July 6th, 2022 at 20:28:50 PDT to describe the Geobacter sulfurreduccens bacteria. Thanks to the murder of local crows, who trained me to feed them unsalted peanuts every morning and every time they saw me riding my bike to campus.

Table of Contents

Table of Figures

1. Introduction

Integration of biological molecules into electronics is an emerging field of study, taking advantage of the bottom-up fabrication approach to create nanoscale devices. These processes can take advantage of the self-assembling nature of biological molecules with the added benefit of greener fabrication methods compared to traditional silicon-based processing techniques.[1] Modeling these systems requires atomic-level simulation, with transition metals adding unique problems due to the variable molecular properties that are dependent on oxidation state and orbital occupation. While purely organic systems are common in biological systems the incorporation of transition metal ions is essential to charge transport in most biological systems, such as use of Fe atoms in cytochromes that participate in cellular redox chemistry or the Mg site in chlorophyll.[2] Applying quantum mechanics models to these systems, other properties come to light including spin-dependent transport enhanced by unpaired spin transport. Using these models, we can better understand the properties of these systems and fabricate a new generation of nanoelectronic devices constructed with these metalated biomolecules.

The approach used to model the nanodevices is comprised of three main steps as shown in Figure 1.1. Structure information can be extracted from experimental data, such as x-ray crystallography or cryo-EM data, or generated using molecular mechanics methods. While molecular dynamics (MD) packages such as AMBER (Assisted Model Building with Energy Refinement) or NAMD (NAnoscale Molecular Dynamics) can be used to model most biological molecules as well as metal ions in solution, these force fields cannot easily be applied to coordinated transition metal atoms. Other methods, such as density functional theory (DFT) or semi-empirical orbital calculation methods, must be used to calculate atomic forces for structure minimization. The coordinates of this final structure can then be used to run a single point DFT calculation to extract molecular orbitals and available energy levels. This Hamiltonian can be used for electronic structure calculations or can be incorporated into Green's functions methods to simulate current flow in a nanodevice with multiple terminals.

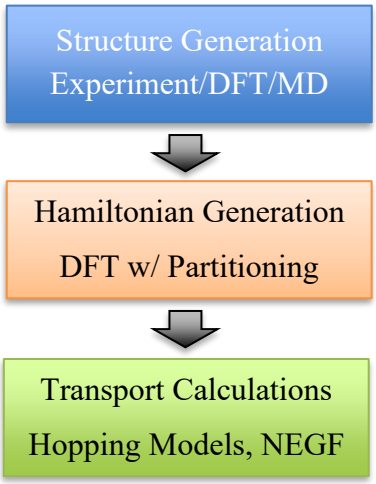


Figure 1.1 - Major steps involved in modeling biomolecules as a nanodevice.

In my studies, I will explore the assumptions made at each step of the modeling process and compare the results between macromolecular systems and nanodevice setups. Validation of the model will come both from the theory used to construct the model and the experimental results observed for the given setups. I will propose new additions to the model to account for effects like

spin-dependent transport and current voltage hysteresis using a combination of empirical approximations and first principles models. The model will be applied to cytochrome proteins and metal modified DNA structures to predict the electronic behavior of these materials. Using these developed models, it is possible to model a biomolecular device with transition metals from the ground up both to explain experimental results and explore new device applications.

a. Metal complexes in proteins

The basic building block of a protein is an amino acid, which varies by organic ligands, and are held together in a polypeptide chain formed by the peptide bond between the amine and carboxylic acid. The protein sequence determines the overall structure and function, with large-scale structure forming based on the peptide sequence. With some exceptions, proteins that are not active in redox chemistry are electrically insulating[3] while electrically active proteins contain transition metals that participate in electron transport. These proteins typically coordinate the transition metal ions and incorporate another organic compound, such as a heme porphyrin ring, as part of the structure. Applying a model that includes the effects of these coordinated transition metals is crucial to understanding the electronic properties of these proteins and using them to construct nanoelectronic devices. Cytochrome proteins are of particular interest due to their use of heme B porphyrins for charge transport in redox reactions. In these systems, the oxidation state of the coordinated Fe center can influence electron transport, which can be modeled in many cases as incoherent electron hopping to facilitate a redox reaction. Other models have also been applied,

such as looking at tunneling with coherent transport or mixing multiple models, described in more detail in the next section. Experimental results have shown that these cytochromes can have high conductivity[4] with temperature dependence that cannot be explained by a simplified transport model.[5]

In this work I have applied the developed model to the OmcS cytochrome, a nanowire chain of six-heme cytochrome monomer units found in the *Geobacter sulfurreducens* bacteria, as well as the small tetraheme cytochrome (STC), a four-heme cytochrome found in the *Shewanella oneidensis* bacteria. Experimental results can be used to validate the results from the model and predict other exciting electron properties, such as spin-dependent charge transport that can be engineered for a new generation of nanodevices. More broadly, creating an accurate model for cytochrome proteins will allow us to predict their behavior when used in nanoscale circuits or used in larger heterostructures as part of a nanodevice.

b. Metal-Modified DNA

There is a lot of interest in DNA heterostructures for use in nanoelectronics due to their self-assembly behavior, stability in solution, and customizability. The DNA structure consists of repeating nucleotides consisting of a nucleobase bonded to a ribose sugar and phosphate group as shown in Figure 1.2. A single strand of DNA (ssDNA) is formed when multiple nucleotides connect via the phosphodiester bond formed between the 5’ carbon of one ribose to the 3’ carbon

of the following ribose. There are four naturally occurring bases that can form two possible base pairs, AT and GC, that form hydrogen bonds to create double stranded DNA (dsDNA). These complementary base pairs on the ssDNA self-assemble into dsDNA in solution, a property that can be used to create large heterostructures known as DNA origami. These highly customizable structures have a wide array of applications including in nanoelectronics, though these applications typically use DNA origami as scaffolds for other molecules due to their low conductivity.[6]

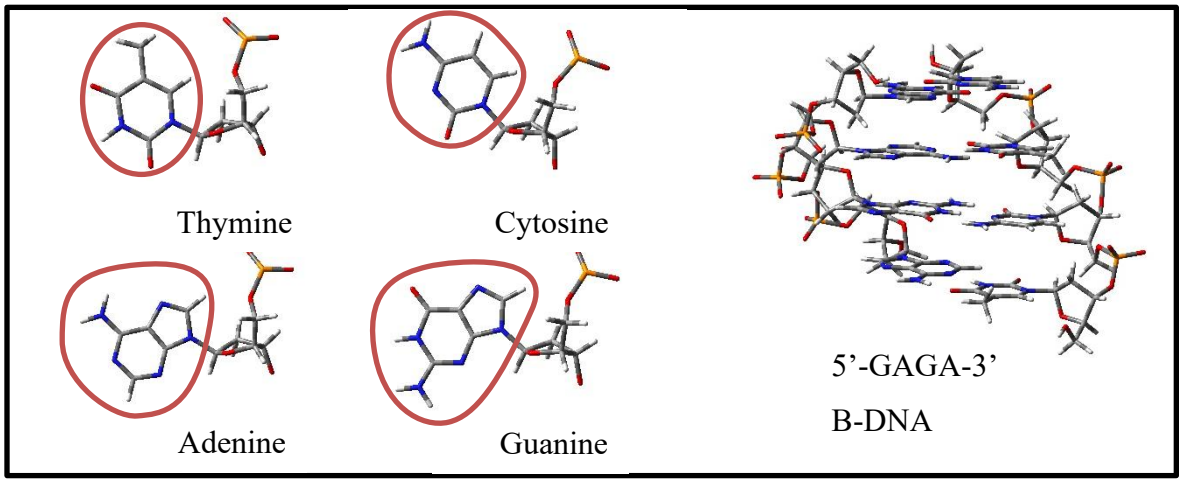


Figure 1.2 - The four standard nucleotides in DNA as well as a four base pair double stranded structure in the B-DNA conformation. The nucleobase is circled in red for each one, with the sugar and phosphate attached.

Modification of the DNA structure has been the subject of a range of studies, ranging from modifying the backbone with addition of ligands to direct modification of the bases themselves. Both methods have been used to add metals to the base pair, either bonded to a ligand outside of the structure[7] or intercalated between the nucleobases.[8] Outside of the four naturally occurring nucleobases, many modified nucleobases have been synthesized that can be incorporated into dsDNA to further enhance customizability.[9,10] The metal-modified DNA (mmDNA) structures can

be formed through a simple experimental setup using a target base pair, such as a cytosine-thymine (C-T) mismatch, and adding a single reagent (such as silver nitrate, AgNO3) as a single annealing step.[8] The resulting mmDNA base pair can have a metal ion located right between the bases, called intercalation, or located beside the DNA basepair, usually one of the two grooves found in the double stranded structure, referred to as the major and minor grooves. The synthesis process has advantages due to specificity of basepair modifications in DNA strand and the simple processing required. This makes mmDNA devices advantageous over larger scale and more intensive silicon-based devices, especially when fabricating complete DNA-based device using the self-assembling DNA origami process. Such a process would allow construction of self-assembling nanodevices with targeted "doping" sites containing metal atoms that can form interconnected circuits within the DNA heterostructure.

c. Literature Review

The models that have been developed for charge transport in transition metals have origins in experiments conducted with metal isotopes in the 1940s followed by detailed spectroscopic analysis of such molecules over the following decades.[11] Initial electron transport models considered discrete electron hopping from a chemistry perspective, looking at reaction kinetics by applying an Arrhenius model. Marcus theory, a hopping model that includes non-adiabatic behavior, has been applied extensively to charge transport in biological systems, especially proteins with transition metals such as cytochromes.[12] Other approaches have been more recently

developed to look at electron transport from a quantum mechanics perspective, looking at coherent wavefunction transport through biomolecules. Both models have merit for describing macromolecular systems, and new developments can be used to apply these models to systems.

Electron Hopping Models

Marcus theory is a semi-classical electron hopping model that uses tunneling magnitude and reorganization energy to determine the kinetics of electron transfer. By constraining electron transport to a discrete transition, the energetics of the system is considered before and after electron transfer, labeling the molecules or sites of interest as donors and acceptors with the following chemical reaction:

$$[D^-] + [A^+] \rightleftharpoons [D] + [A]$$

Marcus theory looks at the shifts in nuclear coordinates during this electron transfer reaction and applies the Frank-Condon principle, which assumes that electron transitions happen much faster than nuclear vibrations. Therefore, electron transfer only occurs during reorganization of the system when the nuclear potential energy of the LHS matches the RHS of the reaction. This means that the free energy barrier (activation energy) for this reaction is determined by the potential energy difference across the nuclear coordinates and the pre-exponential term is determined by the rate of electron transport between the donor acceptor states at the transition point. While it is a challenge to calculate the exact activation energy using atomic simulations, this exponential term

is a function of the total reorganization energy and averaged free energy of oxidation with the rate constant given by

$$K_{et} = A\exp\left(\frac{\Delta G^{\ddagger}}{k_B T}\right) = \frac{2\pi}{\hbar}|H_{ab}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}}\exp\left(-\frac{(\lambda+\Delta G_{ox})^2}{4\lambda k_B T}\right) \quad (1)$$

a function of λ, the reorganization energy, H_{ab}, the charge transfer integral, and ΔG_{ox}, the free energy of oxidation.[13] In this equation, only H_{ab} requires an quantum mechanical orbital calculation, looking at the hopping term from molecular orbitals on the donor and acceptor sites. The free energy of oxidation and reorganization energy are calculated based on calculated total energies, either from quantum mechanics (QM) or molecular mechanics (MM) calculations, along reaction coordinates as shown in Figure 1.3. From this figure, we can calculate the forward reorganization energy as the difference between point A and D, the reverse reorganization energy as the difference between B and C, and the free energy of oxidation as point C minus point D.

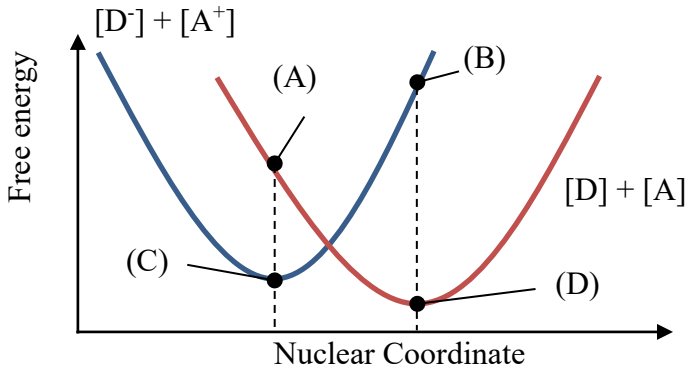


Figure 1.3 - Marcus Theory parabolas showing the state free energy as a function of nuclear coordinates, with the state before electron transition in blue and after transition in orange.

Marcus theory has been used extensively to provide a grounded explanation of many experimental observations, especially redox reactions in proteins. The underlying assumptions explain slow electron transport, where reorganization is the bottleneck to electron hopping and hopping occurs instantaneously relative to this process. Marcus theory has been applied to explain charge transport in several cytochrome systems recently including cytochrome STC in *S. oneidensis*[14] as well as the OmsC cytochrome in *G. sulfurreducens.*[5] Because of the exponential relationship to reorganization energy and free energy of oxidation, the accuracy of these energy calculations is paramount and can impact the electron transfer rate by orders of magnitude.

Quantum Transport with Green's Functions

As an alternative to the classical hopping-based model, a quantum-based model can be constructed based on the interactions of electron wavefunctions. The NEGF method for modeling quantum transport of electrons was constructed to use this approach to fully describe atomic-scale electron transport in nanodevices.[15] In the coherent limit, such a model can describe ballistic transport of electrons in a static, adiabatic macromolecule between electrical contacts. The Landauer-Büttiker formalism is the most prominent approach used to calculate the coherent transport and has been applied in this form to DNA systems[16,17] with results matching up with experimental observations. Coherent transport has also been applied to cytochrome protein systems[18,19] though only over very short length scales. Instead, for larger molecules the NEGF coherent model has been applied in parallel[20] or in series[21] with hopping models because of the

projected loss of coherence between each heme site. The highly fluctuating behavior of macromolecules in biological systems suggests that a coherent model is only accurate for short length or time scales, while the incoherent (hopping) model incorporates non-adiabatic behavior during change transfer including interaction with the surrounding environment.

As an alternative to mixing incoherent and coherent models, there has been significant work on modifying the coherent NEGF model to include interactions with the environment, which can cause dynamic energy level shifts and dephasing of electrons. As a ground-up approach, several groups have modeled quantum transport with phonon interactions directly not only in inorganic systems, such as quantum dots[22] and carbon nanotubes[23] but also DNA systems.[24] However, in these cases the models are applied to simplified systems using approximate Hamiltonians to model nearest neighbor hopping due to the complexity of these models and resulting computational requirements. While using approximate Hamiltonians can lead to accurate structural predictions for a variety of larger biomolecules,[25] the resulting energy level structure can lose accuracy due to lack of consideration of the electron interactions.[26] Instead of applying complex models, the Büttiker/decoherence probe approach uses fictitious probes with self-energies to model decoherence, only adding only a single term to the NEGF formulation, allowing one to use the full DFT Hamiltonian without significantly increased computation. This model has been used to model decoherence in many systems based on experimental data and applied in many forms to predict electron behavior of DNA heterostructures.[17,27,28]

Spin-dependent Transport

Theoretical studies of spin-dependent quantum transport were published soon after the first observations of molecular conductance in break junctions more than 20 years ago.[29,30] These studies look at total electron transmission through small organic molecules with ferromagnetic contacts, focusing specifically on the dependence of transport on the magnetic polarization of the contacts and the contact coupling.[31,32] Experimental results confirm some of these findings but show dependence on the environment not explained by these models.[33] More recent models have included corrections that account for spin-lattice interactions and spin-orbit interactions to enhance the model use for electrical contacts,[34] but these models still fail to explain the mechanisms of many observed spin-dependent effects such as chiral-induced spin selectivity.[35] Spin dependent transport has been observed in biomolecules, such as DNA, because of such effects.[36]

The addition of inorganic transition metals adds further complexity to these systems due to the presence of d-orbitals with partial occupation, resulting in electronic configurations that depend on the element and oxidation state. Such configurations have been predicted at the Fe center found in heme[37] which can be found in cytochromes. There have been limited studies modeling coherent spin-dependent transport in cytochromes with some success, although the mechanism of the reported results are not clear.[19]

2. Theory and Methods

To characterize the electronic properties of the biomolecules, a procedure was followed building on previously developed transport models.[38] First, a characteristic structure was generated based on experimental or modeling methods, followed by ab-initio DFT calculations to generate the Hamiltonian for the full atomic structure. This Hamiltonian was then used to study electron transport properties, either looking at the energies and tunneling rates between local orbitals in a hopping model or using NEGF methods to calculate quantum transmission of electrons in the system. The quantum transmission model was extended to include spin-dependent effects found in systems with coordinated transition metals.

a. Structure generation

Investigation of a biomolecule's electronic transport properties begins with a set of atomic coordinates, which establish the basis for constructing electronic orbitals, calculating the electron density, and deriving the Hamiltonian required for quantum transport calculations. The goal of this structure is to be time-averaged to represent the atomic positions while the molecule is being used as a nanodevice. By itself this can be an unrealistic picture of the system since biomolecules are inherently quite floppy, but this can be corrected by adding environment interaction terms in the transport model. With the variety of biomolecules, several different methods can be used for structure determination. As with all modeling techniques, for a given molecule it is best to choose

the method that is most computationally efficient and most closely follows the experimental data available. As a general practice, all molecular structures are determined based on experimental data or residue information for similar structures as a starting point, with modeling tools applied to refine the model as needed.

X-ray crystallography has been employed for several decades to resolve the atomic structure of biomolecules by analyzing the diffraction pattern of X-rays scattered by a crystalline lattice array of the biomolecule, thereby reconstructing the spatial arrangement of atoms. Techniques for x-ray crystallography have improved significantly over this time, with resolutions less than 0.5 Å reached for smaller proteins.[39] More recently, single-particle cryogenic electron microscopy (cryoEM) has emerged as a method for determining macromolecular structure without the need for crystallization. Winning the Nobel Prize in Chemistry in 2017, the technique can generate structures with a resolution of 3-4Å even for larger proteins. Both methods generate 3D maps of electron density, which can be fitted with atomic structures to balance predicted atomic location with physical bonding parameters. While the resulting structure is directly extracted from the molecule itself, the fitting methods and resolution of the structure can cause structures that may lead to unrealistic electronic properties, especially if electronic orbitals are constrained by bonding or inter-atomic distances. In addition, the structure is extracted from a crystal structure or at cryogenic temperatures, which may not match the environment of the nanodevice being modeled.

Experimental methods are not available for all biomolecules, especially for proposed structures where synthesis or characterization techniques are not yet known. While bonding data (bond length, bond angle, dihedral angle, partial charge, hydrogen bonding, and so on) can be applied to structures based on rules of thumb or using data from similar structures, MM methods have been developed for decades to determine atomic structure and look at fluctuations over time. Molecular dynamics (MD) uses empirical force fields calculated between atoms and solves Newton's equations of motion to calculate atomic trajectories. Given a vector of atomic coordinates given by $\boldsymbol{r} = (r_1, \ldots, r_n)$, the equations of motion for the system are

$$F_i(\boldsymbol{r}) = \frac{\partial E_i(\boldsymbol{r})}{\partial r_i} = \frac{\partial}{\partial r_i} \sum_j E_{ij}\left(r_i, r_j\right), \quad \frac{\partial^2 r_i}{\partial t^2} = \frac{F_i(\boldsymbol{r})}{m_i} \tag{2}$$

where $F_i(\boldsymbol{r})$ is the total force on atom *i* and $E_{ij}\left(r_i, r_j\right)$ is the potential between atom *i* and atom *j*.Atomic trajectories are calculated by integrating the equations of motion for each atom, usually using the Verlet algorithm with a fixed time step.[40] Numerous force fields have been developed for different types of materials and molecules, with several different force field parameter sets used for biological molecules.[41] The AMBER force field used in this study parameterizes the total potential energy (E) of the system as the sum of bonded and non-bonded interactions. The total energy is expressed as

$$E_{\text{total}} = E_{\text{bonded}} + E_{\text{non-bonded}} \tag{3}$$

The bonded energy term is given by

$$E_{\text{bonded}} = \sum_{\text{bonds}} k_b(r - r_0)^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] \qquad (4)$$

where r, θ, and ϕ are the bond length, bond angle, and dihedral angle, k_b and k_θ are the bond and angle force constants, r_0 and θ_0 are the equilibrium bond lengths and angles, and n, V_n, and γ define the dihedral parameters. The non-bonded contribution is given by

$$E_{\text{non-bonded}} = \sum_{i<j} \left[\sqrt{\epsilon_i \epsilon_j} \left(\left(\frac{R_{min,i} + R_{min,j}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,i} + R_{min,j}}{r_{ij}} \right)^{6} \right) + \frac{q_i q_j}{\varepsilon r_{ij}} \right] \qquad (5)$$

where distances between atoms are given by r_{ij}, Lennard Jones parameters are given by the van der Waals radius $R_{min,i}$ and the binding energy ϵ_i for each atom, and Coulomb interactions are modeled with atomic charges q_i and dielectric constant ε. Using MD one can generate a time-averaged structure to compute transport properties, either by force minimization across the structure or clustering techniques of a structure along the MD simulation trajectory. Both techniques have been used to generate protein[42] and DNA[27] structures for transport calculations.

Force fields may not exist for every atom in a proposed molecule, especially for coordinated transition metals that may be introduced to an existing biomolecule. While new force fields can be generated based on the interaction of the metal atoms with nearby atoms, fitting to a force field may lead to loss of information about system dependent electronic behavior, especially partial charge and metal-ligand interactions away from equilibrium. For this reason, applying quantum

mechanical (QM) methods can be necessary to compute forces. Using multiple single point DFT calculations (and looking at the perturbation response for each atom), atomic force can be calculated from the energy changes. Minimization of the DFT energy by shifting atomic positions can be used to find the energy-optimized structure. This may be a good reference structure for electronic structure calculations, but further ab-initio molecular dynamics studies, calculating the QM forces at each time step, can also be used to look at structure fluctuations over time. While these methods can be computationally intensive, especially in the case of full ab-initio molecular dynamics, it is possible to selectively apply QM methods. Partial optimization of lighter atoms (like hydrogen) in a structure can be used to approximate a ground state structure. In addition, hybrid QM/MM approaches have been used frequently to investigate chemical reactions and look at structural changes in biomolecules[43] considering the QM and MM layers of interactions to calculate energy differences. As a final approach, semi-empirical methods like PM6 can be used to minimize (or partially minimize) structure with a high degree of accuracy.[44] In our work, these methods have been applied based on the availability of crystal structures, force fields, and computational limits.

b. Self-Consistent Field (SCF) Methods

Determining the electronic structure of a material using quantum mechanics is a complex many-body problem that has been historically very difficult to solve. In fact, it is impossible to find an exact analytical solution to the Schrödinger equation for a system as simple as the helium

atom, due to the many interacting degrees of freedom.[45] Applying the Born-Oppenhemer approximation, we can separate the behavior of the nuclei separate from the electrons to get the following expression for the Schrödinger equation:

$$\widehat{H}\Psi(\boldsymbol{r_1}, \dots, \boldsymbol{r_N}) = \left[\sum_{i=0}^{N}\left(\frac{-\hbar^2\nabla^2}{2m_i^2} + V(\boldsymbol{r_i})\right) + \sum_{i<j}^{N} U(\boldsymbol{r_i}, \boldsymbol{r_j})\right]\Psi(\boldsymbol{r_1}, \dots, \boldsymbol{r_N})$$

$$= E\Psi(\boldsymbol{r_1}, \dots, \boldsymbol{r_N}) \quad (6)$$

for an N electron system, given by the multi-electron wavefunction $\Psi(\boldsymbol{r_1}, \dots, \boldsymbol{r_N})$, where $V(\boldsymbol{r_i})$ is the potential field contribution from the nuclei and $U(\boldsymbol{r_i}, \boldsymbol{r_j})$ is the electron-electron interaction energy. To simplify the problem further approximations are typically used including removing relativistic effects or interactions, using a limited number of basis functions to constrain the system (such as using a grid of k-points in periodic systems or an approximate subset of orbitals in molecular systems), and applying a mean-field approximation. Self-consistent field methods describe using trial wave function with the Schrödinger equation (Equation 6) iteratively, minimizing the total energy to determine the ground state. This minimization occurs on the energy functional, which has the wavefunction as an input, forming the cycle of field equations that must be solved to generate electronic structures.

The Hartree-Fock (HF) method was the first of such self-consistent field methods,[46] using a single Slater determinant to represent the wavefunction of all electrons given by

$$\Psi(\boldsymbol{r_1}, \ldots, \boldsymbol{r_N}) = \frac{1}{\sqrt{\mathrm{N}}} \begin{vmatrix} \phi_1(\boldsymbol{r_1}) & \phi_2(\boldsymbol{r_1}) & \ldots & \phi_n(\boldsymbol{r_1}) \\ \phi_1(\boldsymbol{r_2}) & \phi_2(\boldsymbol{r_2}) & \ldots & \phi_n(\boldsymbol{r_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{r_n}) & \phi_2(\boldsymbol{r_n}) & \ldots & \phi_n(\boldsymbol{r_n}) \end{vmatrix} \tag{7}$$

with normalized single-coordinate wavefunctions given by $\phi_i(\boldsymbol{r_j})$ and the many-body wavefunction allowing for antisymmetric exchange of electrons according to the Pauli exclusion principle. Applying the variational principle to Equation 6 given these constraints gives us the single-electron operator given by

$$\hat{F} = \hat{H} + 2\hat{J} - \hat{K} \tag{8}$$

also called the Fock operator with $\hat{H}$ representing the one-electron operator given by

$$\hat{H} = \frac{-\hbar^2\nabla^2}{2m_i^2} + \sum_{i=0}^{N} \frac{Z_i}{r_{1i}} \tag{9}$$

with the distance between atom *i* and the integral coordinate represented by r_{1i}. The two-electron operators are given by

$$\begin{aligned} \langle \psi_i | \hat{J} | \psi_j \rangle &= \sum_{k,l} \rho_{kl} \int \int \psi_i^*(\boldsymbol{r_1}) \psi_k^*(\boldsymbol{r_2}) \frac{1}{|\boldsymbol{r_1} - \boldsymbol{r_2}|} \psi_j(\boldsymbol{r_1}) \psi_l(\boldsymbol{r_2})\, d\boldsymbol{r_1}\, d\boldsymbol{r_2} \\ \langle \psi_i | \hat{K} | \psi_j \rangle &= \sum_{k,l} \rho_{kl} \int \int \psi_i^*(\boldsymbol{r_1}) \psi_k^*(\boldsymbol{r_2}) \frac{1}{|\boldsymbol{r_1} - \boldsymbol{r_2}|} \psi_l(\boldsymbol{r_1}) \psi_j(\boldsymbol{r_2})\, d\boldsymbol{r_1}\, d\boldsymbol{r_2} \end{aligned} \tag{10}$$

which are the coulomb and exchange operator respectively with density matrix elements given by ρ_{kl}. In fact, each of these operators is typically solved as a matrix across finite basis given by ψ_i. Minimizing the energy of the system gives us the eigenvalue problem

$$\left(\hat{F} - \epsilon\hat{S}\right)\hat{C} = 0 \tag{11}$$

where $\hat{S}$ is the overlap matrix (where the deviation from the identity matrix represents the nonorthogonality of the basis set), ϵ are the energy levels (eigenvalues), and $\hat{C}$ are the coefficients of the molecular orbitals (eigenvectors). These derived orbitals are then used to populate the density matrix as part of the SCF. While this exact exchange term can give relatively accurate results, it does not include electron correlation energies and the calculations do not scale well to large systems.

As an alternative, density functional theory (DFT) builds on the mean-field theorem to use an energy functional that is based on the electron density $\rho(\boldsymbol{r})$. Instead of an exact exchange term, an exchange-correlation functional is used, and the total energy is minimized with the variational principle. This leads to the Kohn-Sham equations that can be used to solve for molecular orbitals, given by

$$\left[-\frac{\hbar^2\nabla^2}{2m} + V_{ext}(\boldsymbol{r}) + \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3\boldsymbol{r}' + V_{xc}[\rho(\boldsymbol{r})]\right]\psi_i(\boldsymbol{r}) = E_i\psi_i(\boldsymbol{r}) \tag{12}$$

where $V_{xc}[\rho(\boldsymbol{r})]$ is the exchange-correlation functional and $V_{ext}(\boldsymbol{r})$ is the potential given by the nuclei and external fields. This equation can be similarly solved as the eigenvalue problem in Equation 11, with the Fock matrix including the exchange-correlational functional instead of the exchange term. In the simplest case the local density approximation (LDA) can be used to generate an exchange-correlation functional, but other functionals include the gradient of the electron density, called the generalized gradient approximation (GGA). Hybrid functionals that include a Hartree-Fock term have emerged as a more accurate functional for describing a wide range of

molecular properties, such as the B3LYP (Becke, 3-parameter, Lee–Yang–Parr) exchange-correlation functional which includes GGA, LDA, and HF terms and is used in our studies.[47]

c. Marcus Theory

To apply Marcus theory to calculate the rate constants given by Equation 1 ab-initio methods were used to determine values for λ, the reorganization energy and H_{ab}, the charge transfer integral. This replaces molecular dynamics methods typically used for such calculations, instead using the SCF energies and orbitals for the ground states of the system as calculated by DFT. The reorganization energy was calculated by taking differences between the minimized DFT energies of the combined system, which create representative geometries that represent the minima of the Marcus theory parabolas. To simplify the computational complexity of these calculations, rather than applying minimization to full donor and acceptor pair, only the single half reaction was considered. This reaction is given by

$$[D]^{\cdot -} \rightleftharpoons [D] + e^{-} \tag{13}$$

where $[D]$ represents the oxidized case (acceptor) and $[D]^{\cdot -}$ represents the reduced case (donor). Because each donor/acceptor site is modeled using this characteristic reorganization energy, the free energy of oxidation, ΔG_{ox}, was set to zero for rate calculation. While this is an approximation because it doesn't necessarily include the coupling effects from surrounding groups, it eliminates the need to repeat the calculation across multiple pairs of donor and acceptor sites, providing a

single characteristic reorganization energy for comparison as well as an analysis of the electronic structure changes during coordinate shifts.

The charge transfer integral, H_{ab}, was calculated using a local orbital basis to measure the hopping energy between neighboring donor and acceptor sites. To capture this, the DFT calculations are applied to the fully interacting system, which allows us to use linear transformations to partition the Hamiltonian by fragmenting and extracting off-diagonal orbital hopping terms. First, separated blocks of the Hamiltonian were diagonalized, and constructing a transformation matrix for the combined Hamiltonian given by

$$\mathrm{V} = \begin{bmatrix} V_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & V_i \end{bmatrix} \tag{14}$$

where V_i is a square matrix of eigenvectors for block i. The transformed Hamiltonian is given by

$$H' = V^{\dagger} H\, V \tag{15}$$

where rows and columns represent the local molecular orbitals for each block and off-diagonal terms represent tunneling energies between molecular orbitals. Note that an orthonormal basis must be used for this Hamiltonian, which can be constructed from using a Lowdin transform from the single-electron Fock matrix and overlap matrix:

$$H = \hat{S}^{\frac{1}{2}} \hat{F} \hat{S}^{\frac{1}{2}} \tag{16}$$

Charge transfer integrals can be calculated by averaging tunneling energies from the highest occupied molecular orbitals (HOMOs) of the donor fragment to thermodynamically available orbitals ($\Delta E \leq k_B T$) in the acceptor fragment at the edge of the band gap. To look at the effects of charging and reorganization, especially when looking at electron hopping in systems, one must consider the charge transfer integral between an occupied state of a donor site and the unoccupied state of an adjacent acceptor site. In these cases, the charge transfer integral can be calculated from the HOMOs of the donor to the lowest unoccupied molecular orbital (LUMO) of the acceptor (looking at N multiple LUMOs if $\left|E_{LUMO} - E_{(LUMO+N)}\right| \leq k_B T$). As a comparison, charge transfer can be calculated from the HOMOs of the donor fragment to the HOMOs of the acceptor fragment in the donor state, including all thermally available HOMOs in both fragments ($\left|E_{HOMO} - E_{(HOMO-N)}\right| \leq k_B T$). Comparing the HOMO-HOMO and HOMO-LUMO couplings can elucidate the effects of charging (electronic orbital re-arrangement) during electron tunneling or hopping events. In all calculations, hopping energies must be calculated between orbitals of the same spin (alpha or beta) except in the singlet case when alpha and beta orbitals were always found to be equivalent.

d. Spin-dependent Model Theory

An accurate model of spin state is necessary when modeling electronic orbitals in any molecule. In the case of organic structures with light atoms and even numbers of electrons, it is typically assumed that spin state energy splitting is minimal, and all electronic orbitals contain

electron pairs. However, there are cases where spin states can split significantly, and additional degrees of freedom must be considered during quantum mechanics calculations. For transition metals specifically, the oxidation state and spin properties are dependent on the geometry and symmetries of the inorganic system, where the partially occupied d and f shells can significantly change the electrical properties of the system. Before considering this, let us first derive the energy expression for a system where all electrons are paired and electronic spin is not a variable in the calculation, referred to as the restricted closed-shell case. For Hartree Fock (HF) calculations, the total energy for the restricted closed-shell case is given by

$$E = tr\left[\rho \times \left(\hat{H} + \hat{J} - \frac{1}{2}\hat{K}\right)\right] \tag{17}$$

with one pair of electrons per orbital.[48] Note these electron integrals can be expanded to the DFT case by replacing the J and K matrices with the Coulomb and exchange-correlation operators that are density functionals, but we will only consider the HF formulation for this section. The Fock operator can be diagonalized to give the single-electron energies of the system, given by ϵ in Equation 11, that obey Koopman's theorem, meaning that occupied energy levels correspond to ionization energies and unoccupied energy levels correspond to electron affinities.[49] For the cases of unpaired spin it is necessary to treat spins separately, where the unrestricted open-shell case is constructed by splitting the Fock operator into two arbitrary spin directions, labeled as alpha and beta spin. This gives us two Fock operators that in the HF formulation are written as

$$\hat{F}_\alpha = \hat{H}_\alpha + \hat{J}_\alpha + \hat{J}_\beta - \hat{K}_\alpha$$
$$\hat{F}_\beta = \hat{H}_\beta + \hat{J}_\alpha + \hat{J}_\beta - \hat{K}_\beta \quad (18)$$

which resolves to Equation 8 when the alpha and beta spins are equivalent. This formulation allows us to properly construct various spin states by modifying the number of alpha and beta electrons, whose total energy can be compared to determine the lowest energy spin configuration. However, these orbitals are not constrained to be eigenfunctions of the total spin operator, meaning that the spatial parts of alpha and beta orbitals may not completely overlap, due in part to the fact that density matrices are constructed for each spin independently. Spin contamination arises from the loss of the total spin quantum number due to the mixing of alpha and beta spin states, with the specific formula given by

$$\delta_s = N_\beta - \sum_{i\in\alpha}\sum_{j\in\beta} |\langle i|j\rangle|^2 \quad (19)$$

In the limit of equal numbers of alpha and beta occupied states, the self-consistent field (SCF) can converge to the RHF formulation, with alpha and beta orbitals matching exactly and therefore no spin contamination. Applying additional constraints to the orbital sub-spaces can ensure that the total spin quantum number is retained, which is the framework for the restricted open-shell model. In this formulation, rather than separating the molecular orbitals into alpha and beta spin, they are separated into doubly-occupied, partially-occupied, and valence molecular orbitals.[50] Using operators for each shell, the equation for the total energy is given by

$$E = 2tr\left[\hat{\rho}_c \times \left(\hat{H} + \hat{J}_c - \frac{1}{2}\hat{K}_c\right)\right] + tr\left[\hat{\rho}_o \times \left(\hat{H} + 2\hat{J}_c - \hat{K}_c + \frac{1}{2}\hat{J}_o - \frac{1}{2}\hat{K}_o\right)\right] \quad (20)$$

which accounts for single electron energies as well as all interaction terms. Unique Fock matrices can be constructed for the closed, open, and valence shells with the following definitions

$$\begin{aligned} \hat{F}_c &= \hat{H} + 2\hat{J}_c - \hat{K}_c + \hat{J}_o - \tfrac{1}{2}\hat{K}_o \\ \hat{F}_o &= \hat{H} + 2\hat{J}_c - \hat{K}_c + \hat{J}_o - \hat{K}_o \\ \hat{F}_v &= 2\hat{F}_c - \hat{F}_o = \hat{H} + 2\hat{J}_c - \hat{K}_c + \hat{J}_o \end{aligned} \tag{21}$$

with the added constraint that doubly occupied orbitals (F_c) must be filled before singly occupied (F_o) orbitals to generate the corresponding density matrices.[51]

From a quantum chemistry perspective, if we do not consider spin-orbit coupling or other relativistic terms, the spin quantum number must be retained, meaning that paired closed-shell electrons must be constrained to the same molecular orbital. While the restricted-open formulation ensures this by partitioning orbitals by occupation, problems arise when attempting to generate unique molecular orbitals within each subspace. In the unrestricted case, alpha and beta Fock operators can be treated completely independently for diagonalization, generating unique alpha and beta molecular orbitals to represent the occupied and valence states in the system. This means that it is trivial to form a single Hamiltonian operator combining both the alpha and beta Fock matrices to get

$$\hat{F}_{unrestricted} = \begin{bmatrix} \hat{F}_\alpha & 0 \\ 0 & \hat{F}_\beta \end{bmatrix} \tag{22}$$

with off-diagonal terms set to zero signifying spin-spin and spin-orbit coupling are ignored. In the restricted open case, however, orbital occupation is interdependent between shells which leads to

a more complicated formula for the combined Hamiltonian. This operator, called the unified coupling operator, can be formed with complementary-to-projection operators (ρ') assigning relative weights (given by variables a, b, and c) to each matrix to get

$$\hat{H}_{RO} = a\hat{\rho}_c'\hat{F}_c\hat{\rho}_c' + b\hat{\rho}_o'\hat{F}_o\hat{\rho}_o' + c\hat{\rho}_v'\hat{F}_v\hat{\rho}_v' \tag{23}$$

where the subscripts 'c', 'o', and 'v' refer to double-occupied, single-occupied, and valence shells and $\hat{\rho}_i'$ is the projection operator complementary to the shell i and is given by

$$\hat{\rho}_i' = \hat{I} - \hat{\rho}_i \tag{24}$$

where $\hat{I}$ is the identity matrix.[52] The density matrices $\hat{\rho}_i$ represent the occupation within each of the subspaces of the three shells and sum to unity. The weights a, b, and c represent a rotation between each of the subspaces to create a set of canonical orbitals, meaning that the total energy of the multi-electron system remains constant. Therefore, the single electron energies can vary arbitrarily and, as a result, the unified coupling operator does not follow Koopman's theorem.

As an alternative approach, the restricted open method can be written in terms of the unrestricted alpha and beta matrices, applying constraints to the orbitals to eliminate spin contamination while giving separate alpha and beta Fock matrices. These can be written as a function of the restricted open-shell doubly occupied and singly occupied Fock operators, with the transformation given by

$$\begin{aligned} \hat{F}_\alpha &= \hat{F}_o \\ \hat{F}_\beta &= 2\hat{F}_c - \hat{F}_o \end{aligned} \tag{25}$$

Because these matrices also satisfy the unrestricted open-shell equations (Equation 3), they form a set of unique orbitals whose eigenvalues satisfy Koopman's theorem.[53] In practice, due to the nature of the convergence criteria, the spin contamination of these orbitals is non-zero but greatly reduced from the unrestricted case. The combined Hamiltonian can thus be created using Equation 22 for quantum transmission calculations. This Hamiltonian assumes collinear spin, meaning that the spin operator commutes with the Hamiltonian and there is no mixing of spin states.

Modeling higher order effects, such as spin-orbit coupling and spin-spin interactions within the structure require additional terms, either fully solved within the SCF or as a perturbation to the calculated Hamiltonian. To include these cases, all restrictions on spin direction must be removed and a spinor basis must be defined for each molecular orbital. Building on Equation 22 we can include the off-diagonal terms to get the full generalized open-shell Hamiltonian, $\widehat{H}_G$, given by

$$\widehat{H}_G = \begin{bmatrix} \hat{F}_\alpha & \hat{F}_{\alpha\beta} \\ \hat{F}_{\beta\alpha} & \hat{F}_\beta \end{bmatrix} \tag{26}$$

which includes spin-spin interactions.[54] Including these non-collinear interactions, spin direction must be defined for the α direction, with all other spin directions set accordingly.

e. Green's Function Transport Calculation

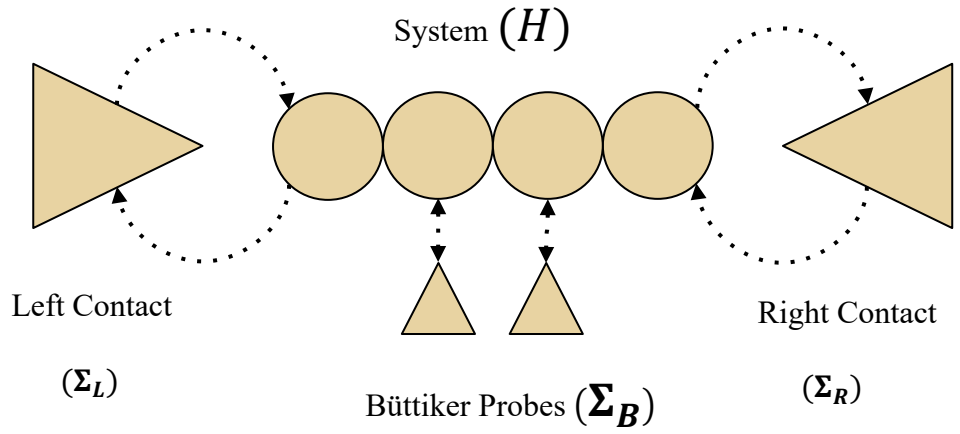


Figure 2.1 - A demonstration of the NEGF method applied to a four-atom system, with a left contact applied to the left-most atom, right contact applied to the right-most atom, and decoherence probes applied to all other atoms.

To measure decoherent transport in a nanodevice, Green's functions were used to model the system and contacts as shown in Figure 2.1. In this Figure, the Σ matrices are diagonal matrices of self-energies, with the left and right contacts having non-zero values for the orbitals in atom 1 and atom 4 respectively. This model allows us to consider the quantum transmission of electrons as a function of energy based on all molecular orbitals of the molecule and including contributions from the device contacts and decoherence (modeled as decoherence probes). The NEGF method was applied to the system Hamiltonian generated from the Fock and Overlap matrices using a Löwdin transformation as defined in equation 16. The transmission spectrum was determined using the Green's function method, starting with the retarded Green's function given by

$$\left[E\hat{I} - \left(\hat{H} + \hat{\Sigma}_L + \hat{\Sigma}_R + \hat{\Sigma}_B\right)\right]\hat{G}^r = \hat{I} \tag{27}$$

where E is the energy, H is the system Hamiltonian (from equation 16), $\hat{\Sigma}_{L/R}$ is the left/right contact self-energy, and $\hat{\Sigma}_B$ is the self-energy of the phase breaking decoherence probes that are used to model decoherence in the system.[28] Note that the retarded Green's function is a matrix that spans the orthonormal orbital basis, as are the self-energy matrices. For each of these self-energy matrices, the wide-band limit approximation is used, which ignores the real part of the matrices and treats them as energy-independent parameters defined at each atom/orbital at the contact.[55] The left and right contact energies are given by $\Sigma_{L/R} = \frac{-i\Gamma_{L/R}}{2}$ for diagonal matrix elements at orbitals connected to the left and right contacts and the decoherence probe self-energies are defined similarly as $\Sigma_{B.n} = \frac{-i\Gamma_n}{2}$ for diagonal matrix elements corresponding to each of the n probes. The Green's function can be directly used to calculate the local density of states (LDOS) at an energy, given by

$$LDOS(\psi_i, E) = -\frac{1}{\pi} Im\left[\hat{G}^R(i,i,E)\right] \tag{28}$$

where ψ_i is the i-th basis orbital and $G^R(i,i,E)$ is the diagonal element of the Green's function at that energy. This can be used to look at local density of states for different fragments of the system, summing across orbitals in each molecule fragment. The diagonal components can be summed to give the total density of states (DOS) as a function of energy:

$$DOS(E) = -\frac{1}{\pi} \mathbf{tr}\left[Im\left[\hat{G}^R(E)\right]\right] \tag{29}$$

Rather than having discrete states, which could be represented as delta functions at each energy level, the broadening contribution from the contacts creates a continuous DOS spectrum across energies. A transmission spectrum can also be obtained from the Green's function, and we can define the transmission at energy E between probe n and m to be

$$T_{nm}(E) = \mathbf{tr}\left[\hat{\Gamma}_n \hat{G}^r(E) \hat{\Gamma}_m \left[\hat{G}^r(E)\right]^\dagger\right] \tag{30}$$

where probe n and m are inclusive of electrical contact or Büttiker probes. Plugging this into the Landauer equation gives a general equation for current through a given probe as

$$I_n = \frac{2q}{h} \sum_m \int T_{nm}(E)[f_n(E) - f_m(E)]dE \tag{31}$$

where $f_n(E) = 1 + exp\left(\frac{E - E_{f,n}}{K_b T}\right)$ is the Fermi distribution defined at probe n. Since the total current at each Buttiker probe is zero and current only flows between the left and right contacts, the transmission term can be simplified and combined to give

$$T_{eff} = T_{LR} + \sum_n \sum_m (T_{Ln} W_{nm}^{-1} T_{mR}) \tag{32}$$

where n and m are summed over the decoherence probe sites and W_{nm}^{-1} is the inverse of $W_{nm} = (\sum_{n \neq l} T_{nl})\delta_{nm} - T_{nm}(1 - \delta_{nm})$, an operator that propagates scattering at each probe site. Therefore, the first term in Equation 32 defines the coherent transport while the second accounts for decoherence. From this transmission the zero-bias conductance by taking the derivative of Equation 31 with respect to the chemical potential difference and plugging in equation 32 to get

$$G(E_f) = \frac{2q^2}{h} \int T_{eff}(E) \frac{\partial f(E - E_f)}{\partial E} dE \tag{33}$$

which can be directly compared to experimental results. Note that in the low temperature limit the effective transmission is proportional to the zero-bias conductivity at that energy.

Adding spin to these calculations can be done using the full spin Hamiltonian, $\widehat{H}_G$, defined in equation 26 and building the full overlap matrix given by

$$\hat{S}_{spin} = \begin{pmatrix} \hat{S} & 0 \\ 0 & \hat{S} \end{pmatrix} \tag{34}$$

showing that the basis functions are the same for orthogonal spin directions α and β. For the unrestricted open-shell cases, $\hat{F}_{\alpha\beta} = \hat{F}_{\beta\alpha} = 0$, indicating that spin treatment is fully collinear and the spin α direction has arbitrary spatial direction.

To add non-collinear effects, the unconstrained generalized open-shell model was used, which allowed modeling of spatial spin direction as well spin interactions. This allows us to replace spin directions α and β with set spin up and spin down directions, usually aligned to the z-direction (the default in G16). While ferromagnetic contacts were not explored in this work (Fermi energy was set independent of electron spin), spin-orbit coupling was applied to the contact atoms in the device (Figure 7.2) for the Au case using spin-orbit coupling coefficients applied to the pseudopotential. To model the effect of spin on transmission and electron density, the Green's function was similarly separated into four quadrants, using the operators from equations 26 and 34 to get

$$\hat{G}^R = \left(E\hat{S}_{spin} - \widehat{H}_G - \begin{pmatrix} \hat{\Sigma}_L^{\downarrow\downarrow} & \hat{\Sigma}_L^{\uparrow\downarrow} \\ \hat{\Sigma}_L^{\downarrow\uparrow} & \hat{\Sigma}_L^{\downarrow\downarrow} \end{pmatrix} - \begin{pmatrix} \hat{\Sigma}_R^{\downarrow\downarrow} & \hat{\Sigma}_R^{\uparrow\downarrow} \\ \hat{\Sigma}_R^{\downarrow\uparrow} & \hat{\Sigma}_R^{\downarrow\downarrow} \end{pmatrix} \right)^{-1} = \begin{pmatrix} \hat{G}_{\uparrow\uparrow}^R & \hat{G}_{\uparrow\downarrow}^R \\ \hat{G}_{\downarrow\uparrow}^R & \hat{G}_{\downarrow\downarrow}^R \end{pmatrix} \tag{35}$$

where $\hat{\Sigma}_i^{\uparrow\uparrow}$, $\hat{\Sigma}_i^{\downarrow\uparrow}$, $\hat{\Sigma}_i^{\downarrow\uparrow}$, and $\hat{\Sigma}_i^{\downarrow\downarrow}$ are the spin-dependent self-energy matrices (usually $\hat{\Sigma}_i^{\downarrow\uparrow} = \hat{\Sigma}_i^{\downarrow\uparrow} = 0, \hat{\Sigma}_i^{\uparrow\uparrow} = \hat{\Sigma}_i^{\downarrow\downarrow} = \hat{\Sigma}_i$) for contact $i \in \{L, R\}$ and $\hat{G}_{\uparrow\uparrow}^R$, $\hat{G}_{\uparrow\downarrow}^R$, $\hat{G}_{\downarrow\uparrow}^R$, and $\hat{G}_{\downarrow\downarrow}^R$ are the corresponding four quadrants of the retarded Green's function $\hat{G}^R$. Using these operators, we can calculate four separate transmission values from each quadrant

$$\begin{aligned} T_{\uparrow\uparrow} &= tr\left[\hat{\Gamma}_L^{\uparrow\uparrow}\hat{G}_{\uparrow\uparrow}^R\hat{\Gamma}_R^{\uparrow\uparrow}\left(\hat{G}_{\uparrow\uparrow}^R\right)^\dagger\right] \\ T_{\uparrow\downarrow} &= tr\left[\hat{\Gamma}_L^{\uparrow\uparrow}\hat{G}_{\uparrow\downarrow}^R\hat{\Gamma}_R^{\downarrow\downarrow}\left(\hat{G}_{\uparrow\downarrow}^R\right)^\dagger\right] \\ T_{\downarrow\uparrow} &= tr\left[\hat{\Gamma}_L^{\downarrow\downarrow}\hat{G}_{\downarrow\uparrow}^R\hat{\Gamma}_R^{\uparrow\uparrow}\left(\hat{G}_{\downarrow\uparrow}^R\right)^\dagger\right] \\ T_{\downarrow\downarrow} &= tr\left[\hat{\Gamma}_L^{\downarrow\downarrow}\hat{G}_{\downarrow\downarrow}^R\hat{\Gamma}_R^{\downarrow\downarrow}\left(\hat{G}_{\downarrow\downarrow}^R\right)^\dagger\right] \end{aligned} \tag{36}$$

as derived previously,[56] using the formula for $\hat{\Gamma}$ given by

$$\hat{\Gamma}_i^{\alpha\beta} = i\left(\hat{\Sigma}_i^{\alpha\beta} - \hat{\Sigma}_i^{\alpha\beta\,\dagger}\right) \tag{37}$$

with spin directions $\alpha, \beta \in \{\downarrow, \uparrow\}$. In this formulation, $T_{\downarrow\uparrow}$ and $T_{\uparrow\downarrow}$ represent the transmission of electrons from spin up to spin down and vice versa, which are directly a result of non-collinear spin behavior. In the absence of spin-orbit coupling, the $T_{\uparrow\downarrow}$ and $T_{\downarrow\uparrow}$ are zero.

3. Metal Modified DNA Library Development

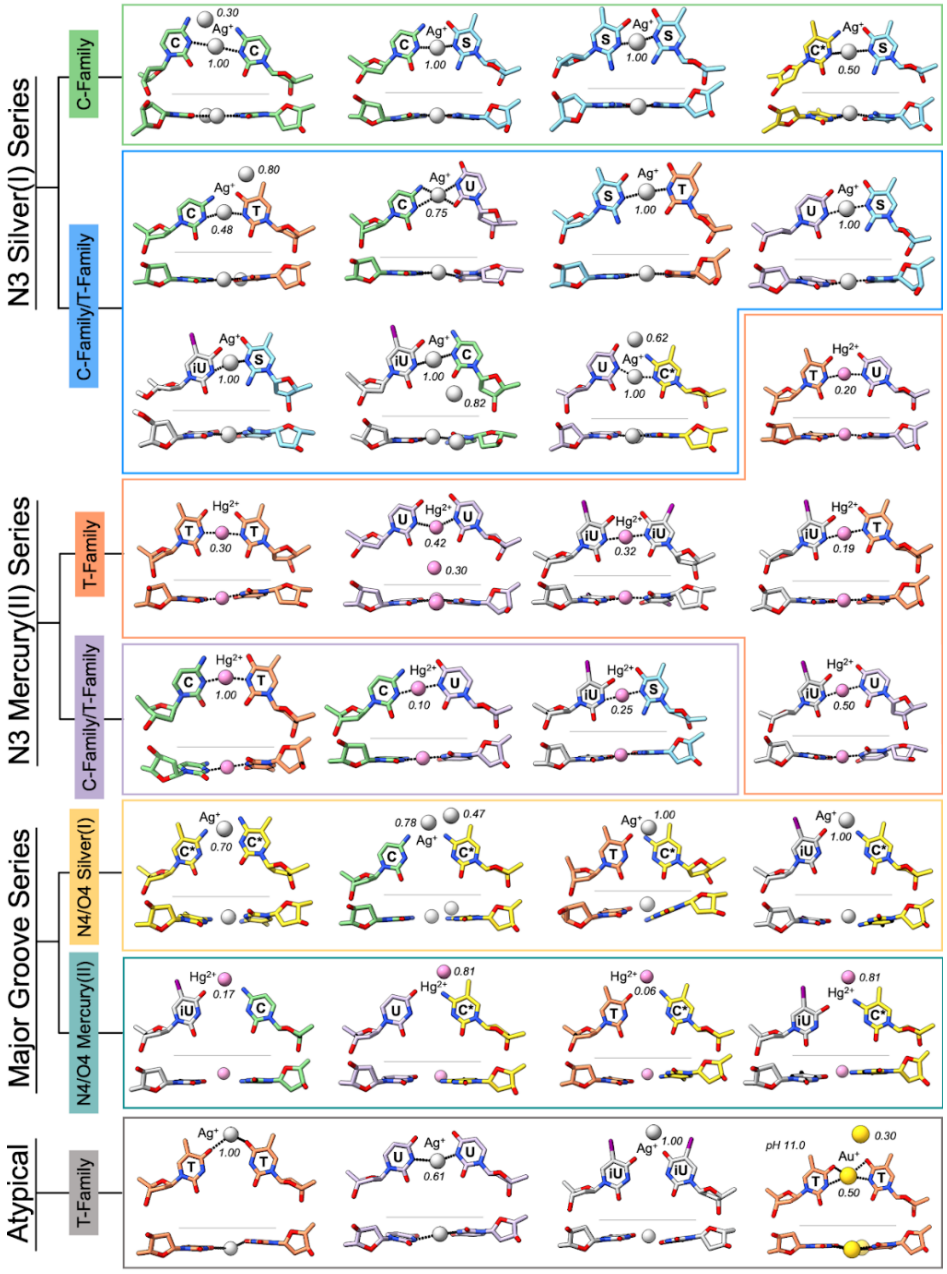


Figure 3.1 - Structural library of all the crystallized mmDNA base pair modeled in this study. Nucleobase carbon atoms are color coded (cytosine (C): green; 5-methyl-isocytosine (S): cyan; 5-methyl-cytosine (C): yellow; thymine (T): orange; uracil (U): lavender; 5-iodo-uracil (iU): white) with heteroatoms displayed (O: red, N: blue). Ag^{+} (grey),*

Hg^{2+} (pink) and Au^{+} ions (gold) are shown with their respective occupancies (italic numbers). Further detailed structural information shown in Figure 3.3 (A) and (B) The chosen base pairs are shown down the helical axis (top image in each pair) and along the radial axis from the viewpoint of the minor groove (bottom image in each pair).

Recently, significant advances have been made in the synthesis and characterization of dsDNA with intercalated transition metals. Upon the discovery of electronic functionality in DNA oligomers containing mmDNA base pairs,[57] many other metal-modified base pairs have been synthesized, a subset of which have been selected for this study and are shown in Figure 3.1. This library of base pairs shows variability in binding modality, either intercalated, in the major groove, or both, categorized by ion identity and base pairing as shown in the left panel. Using the self-assembly property of DNA, the presence of ions was controlled by changing the sequence, making it possible to design materials that exploit this embedded molecular electronics strategy in development of new electronic devices. Combining our models with x-ray crystallography structures derived from experiment, the electronic properties of each base pair could be established and paired to shed light on integration into nanodevices.

We worked with collaborators who synthesized triangular dsDNA origami structures, building on their previous work,[58] which could accommodate metal-modified base pairs using an added linker (see Figure 3.3A). Macroscopic self-assembly was not observed in the absence of metals, demonstrating the requirement, as well as the success, of intercalation of Ag^{+}, Hg^{2+}, and Au^{+} in the trangular structures. X-ray diffraction data from the crystals grown from these structures were used to generate density maps that were used to generate sets of atomic coordinates, referenced

here by their PDB ID. Ab-initio DFT methods were used to minimize the structure and calculate orbitals, followed by Green's function methods to examine electron density across each structure.

a. Modeling Methods

The electronic structure of the various DNA configurations was determined by analyzing the molecular orbitals generated from DFT calculations, with a focus on the HOMO and LUMO energy levels used to calculate the band gap of the molecule. All DFT calculations were conducted on Gaussian 16 software package[59] using a mixed gaussian orbital basis set and the B3LYP hybrid functional. The mixed basis set was used to accurately model different sets of atoms across the periodic table, with the LANL2DZ basis set[60] applied to Ag, Au, and Hg atoms and the 6-31G** Pope basis set[61] applied to all other lighter atoms. A fully paired (no-spin) ground state was determined to be the ground state based on the fully populated d-orbital of the Ag^{+}, Hg^{2+}, and Au^{+} ion and confirmed with a comparison of DFT energies in the triplet state (splitting energy > 10eV). Solvation was applied to the model using the polarizable continuum model,[62] setting the dielectric constant of water to 78.4 (relative to vacuum). Further studies of the local density of states were performed using Green's function methods on the generated Hamiltonian, following the methods outlined previously[27] to generate the two-dimensional DOS plots partitioned by base pair.

While most structure information was extracted from the residue information and x-ray crystallography structure, partial optimization of the metal modified base pairs was used to

generate a stable structure for transport calculations. A seven base pair structure was extracted using 3 base pairs on either side of the modified base pair with OH termination on the 3' and 5' ends (Figure 3.2A). All non-hydrogen atomic coordinates were extracted directly from the PDB structure and hydrogens were added by the GaussView program. For unmodified nucleobases, that is without ligands or metals, all coordinates determined from residue information were unmodified during the optimization process. Due to the nature of the modified base pairs, for instance the halogen (iodine) modified Uracil bases, some manual manipulation of bond lengths and bonding angles were needed to generate a stable structure. As a final step, added hydrogens near the metal atom and their connected groups were optimized to ensure the intercalated metal atoms had minimal force. These partial optimizations were conducted by unfreezing single atoms or ligands and minimizing B3LYP DFT energy of the structure. As a final check, force calculations were conducted for all structures to ensure stability (Figure 3.2B), comparing the maximum forces to that of the baseline metal-free structure.

Force calculations on analyzed crystal structures resulted in atomic forces as high as 5 eV/Angstrom for all atoms, with highest forces occurring on the base pair structure itself rather than the modified ligands or metal atoms. While these calculated forces are relatively high, they are comparable to the modeled forces on the seven base pair segment of an unmodified DNA triangle structure (PDB ID: 5W6W) as well as a standard B-form DNA duplex of the same length.[63]

Density functional theory calculations all converged with a mean-square density difference of less than 1e-8, the default “tight” convergence criterion set for Gaussian16.

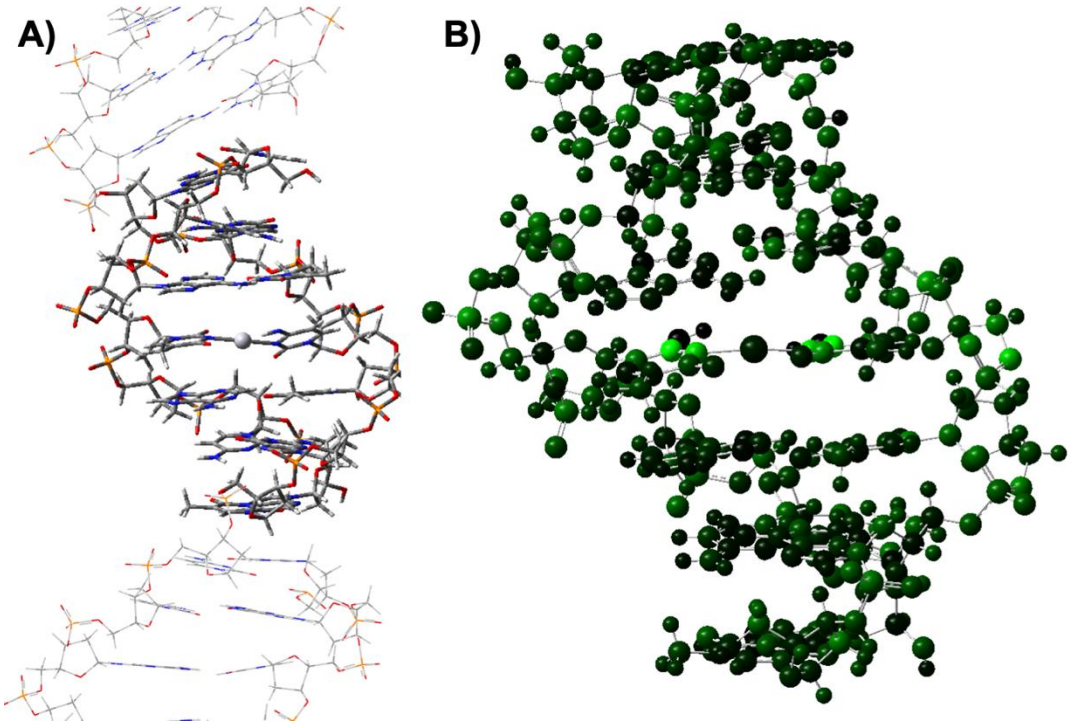


Figure 3.2 - Hg2+-modified DNA structure (dC:Hg2+:dT, PDB ID: 7SD7) used for density functional theory calculations shown in the GaussView software: A) the x-ray crystallography structure with the 7 base pair segment highlighted; B) the isolated 7 base pair segment w with green highlighting to represent atomic forces.

b. Electronic Structure Results

Energy level calculations were carried out on the triangle core using DFT to elucidate an electronic fingerprint of the different classes of structures in our mmDNA library (Fig 5D). We found that all mmDNA pairing modes (N3, N4/O4, Hg2+, Ag+, methylated, halogenated) add levels to the LUMO, akin to the conduction band in a bulk material. Even the addition of a single mmDNA pair causes a distinct lowering of the energy gap in these structures. A comparison of dC:Hg^{2+}:dT, dC:$(Ag^{+})_2$:dT and dC:dG shows similar addition of LUMO states and reduction of the energy gap from 3.90 eV to 3.35-3.42 eV for both metals. Using the extra, unassigned Ag^{+} sites found during dU:Ag^{+}:dS structure refinement, we observed a dramatic decrease in the gap to 2.45 eV for the three-metal base pair case, versus 3.35 and 4.31 eV for the one and zero metal cases, suggesting that harnessing and optimizing this multi-ion effect may lead to further electronic enhancement. A comparison of N3- and N4-bound Ag^{+} and Hg^{2+} carried out by analyzing the methylcytosine analogs to dC:Ag^{+}:dS and dC:Hg^{2+}:dT. Surprisingly, we observed a large reduction in energy gap through a single LUMO state in the major-groove-bound mmDNA structures relative to N3 ones (3.35 eV to 1.86 eV for Hg^{2+}; 3.29 eV to 2.22 eV for Ag^{+}), likely a result of tilting the nucleotides together at the top toward the metal. This offers the possibility of a processing metal chain, rather than the traditional straight, 1D chain of N3 metals that might conduct efficiently along the double helix.[64] Finally, we analyzed dT:Hg^{2+}:dT against similar base pairs with one and two 5-position iodines replacing the thymidine methyl groups. Each iodine added more states to LUMO, dropping the gap from 3.76 to 3.32 to 3.11 eV for zero, one and two

iodines, respectively. We anticipate that increasing the number of mmDNA pairs and the tuning of these electronic effects will allow for the design of electrically-active DNA nanomaterials.

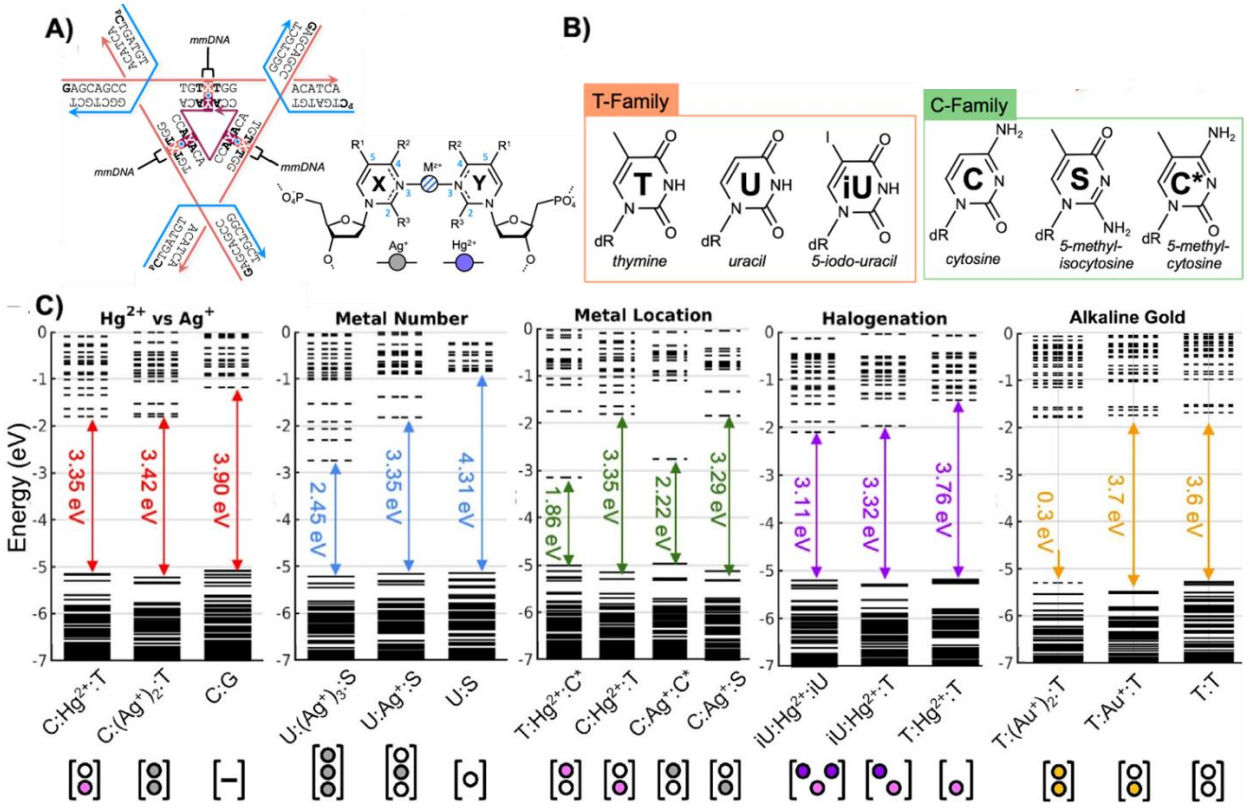


Figure 3.3 – Topology and structure of mmDNA outlined with electronic structure from ab-initio DFT calculations. ***A)*** *The 3D triangular structure used to generate crystals containing mmDNA basespairs, with detail on the modified bases and metal intercalator location.* ***B)*** *The structure of the non-standard bases used to create mmDNA structures.* ***C)*** *Energy gap calculations of the middle seven base pairs of energy-minimized mmDNA crystal structures from our library are carried out using DFT: Hg^{2+} vs. Ag^{+} vs. WC (red); number of metals per base pair from 0-3, using extra sites from the dU:Ag^{+}:dS refinement (blue); N3 vs. major groove N4/O4 metalation (green); and iterative replacement of methyl groups with iodine atoms (purple). Pictoral representations of each base pair are drawn below, with Hg2+ (pink), Ag+ (silver), iodine (purple), unfilled sites (empty circles) and WC pairs (—) indicated. Metal-mediated energy levels are observed in LUMO.*

To investigate the effect of mmDNA base pairs on DNA conductivity, we calculated the DOS as described above on the seven base pair edges of triangle crystal structures with Watson-Crick pairing, Ag^{+} pairing, and Hg^{2+} pairing at the fourth position (see Figure 3.3A and Figure 3.4). The generated DOS spectra demonstrated that the added metal ions created states at the modified base pair with broadening that impacted their surroundings.

The DOS contour plot in

Figure 3.4 shows that, even with just one modified base pair, the density of states within the bandgap increases by several orders of magnitude, increasing the possibility of conduction for the full seven base pair structure. This will increase the expected transmission throughout the structure, showing that the metal modified base pairs can be used to modulate conductivity across a DNA strand for electronics applications.

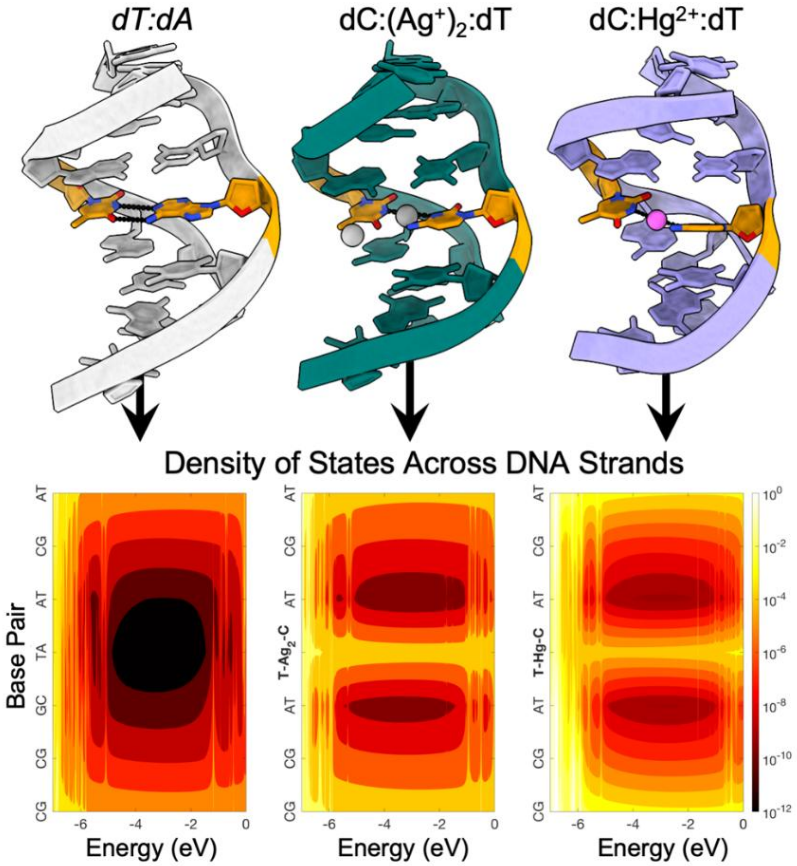


Figure 3.4 - A contour plot showing the DOS spectra across the unmodified and metal modified seven base pair segments, with the y-axis representing the spatial dimension partitioned by base pair and the x-axis representing energy grid, showing DOS as a function of energy. From left to right the structures depicted are the unmodified structure (PDB ID: 5W6W), the Ag^{+}-modified structure (dC:$(Ag^{+})_2$:dT, PDB ID: 7SDS) and the Hg^{2+}-modified DNA structure (dC:Hg^{2+}:dT, PDB ID: 7SD7). Structures are directly shown above each plot.

c. Molecular Orbital Analysis and Results

The molecular orbitals were calculated for dT:Hg^{2+}:dT, diU:Hg^{2+}:dT and diU:Hg^{2+}:diU to determine the effect of halogenation on the density of states in the mmDNA base pair (Figure 3.5). The heavy atoms were seen to be adding energy levels to LUMO, closing the energy gap for the

molecule. For all three base pairs, the lowest LUMO state remains localized to Hg^{2+}, while the LUMO+1,2 states are localized closer to the iodine atoms. From this, it can be inferred that the addition of halogens does not overpower the effect of Hg^{2+}, but rather contributes to the delocalization of the molecular orbitals in the conduction band.

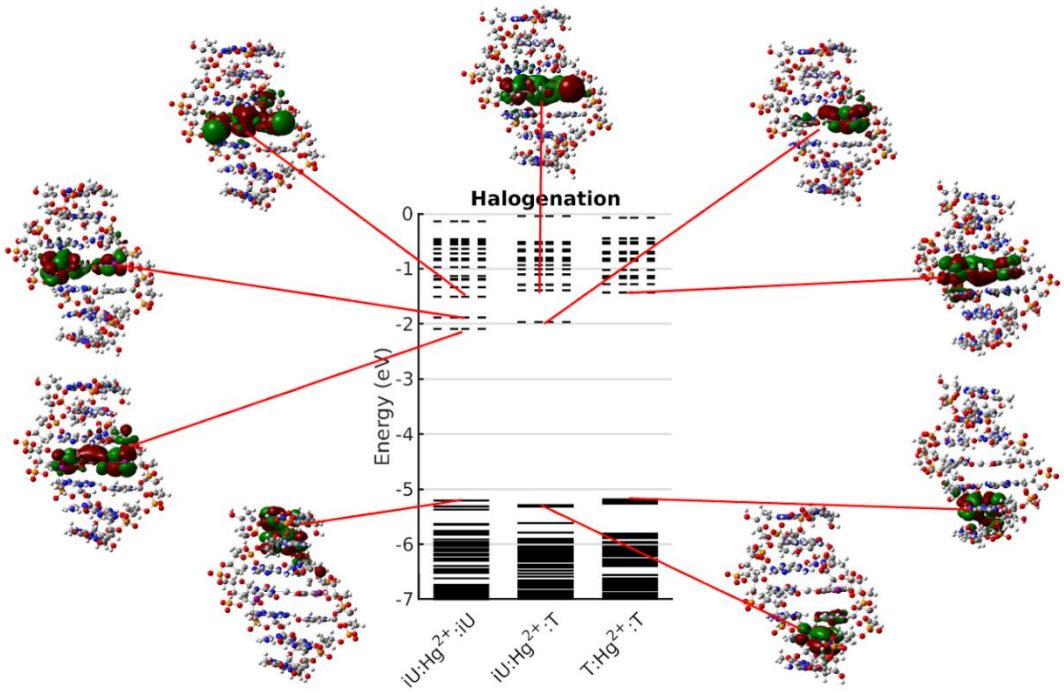


Figure 3.5 - Energy gap determination in Hg^{2+} base pairs with 0,1 and 2 halogens replacing methyl groups at the 5 position of nucleotidyl ring. An isosurface ($|\psi|<0.01$) was generated for the HOMO and LUMO orbitals are calculated for each structure, using red and green surfaces to represent positive and negative orbital lobes. The HOMO appears largely unaffected by the heavy atoms, while LUMO states are added upon addition of mercury and iodine.

Energy gap calculations for the gold base pair, $T:Au^{+}:T$ were carried out with reduced thymine (simulating pH > pKa ≈ 9.5) and oxidized thymine (simulating neutral pH). All thymines in the 7 bp stack were addressed. Protonation of the mmDNA thymines was carried out at the O2 site,

which was calculated to have the lowest overall force, while other thymines were protonated at N3. Three cases were investigated: the mmDNA base pair with both gold sites, the pair with only the N3 Au^{+} site, and the empty mismatch, all derived and energy minimized from PDB structure 7UL8 reported here (Fig. S7). The two-metal pair showed a single LUMO state corresponding to the orbital around the O4 Au^{+}. The localization of this state to the ion suggests that a hopping transport mechanism might be most accessible at this energy level. More states were apparent in the conduction band for the oxidized molecule. Elimination of the O4 Au^{+} removed this LUMO and slightly rearranged the bands. LUMO for these structures was localized to the N3 Au^{+} ion. Energy levels between the structures at this stage were consistent with other modeled mmDNA and unfilled base pairs. pH may play a moderate role in energy gap formation.

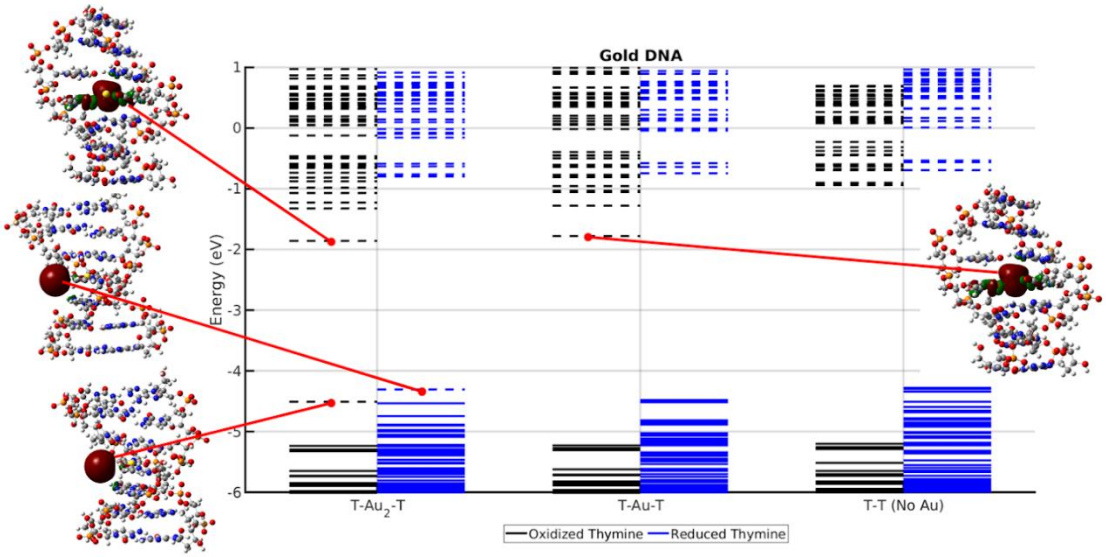


Figure 3.6 - Energy gap calculation and select molecular orbitals for reduced (~ pH 11.0) and oxidized (~ pH 8.0) T:Au^{+}:T base pairs (original structure from PDB ID: 7UL8). An isosurface ($|\psi|<0.01$) was generated for the LUMO

orbitals are calculated for each structure, using red and green surfaces to represent positive and negative orbital lobes.

d. Discussion

DNA is a choice material for construction of a new generation of nanoelectronics created from the ground up, using self-assembled heterostructures that can be programmed with nucleotides. But the high band gap and resulting low conductivity poses a challenge for DNA-based electronics, such as interconnects between each DNA circuit component.[65] Incorporation of transition metals as intercalators in modified base pairs opens up a new set of possibilities for modifying the electronic properties of the device, lowering the band gap and creating a localized charge carrier site that can increase electron transport. The use of crystallography data to generate the DNA structure was crucial for determination of the location of metal binding, which have yet to be included in molecular mechanics models for nucleic acids. Therefore, partial DFT minimization was used to minimize the structure, ensuring any large inaccuracies in structure due to imaging resolution were resolved. Future work combining QM methods to model interactions within the mmDNA base pairs, combined with MM methods for modeling the DNA strand, will be crucial for looking at the dynamics of the mmDNA system in an aqueous environment, especially for helicity information and conformation in an unconstrained system.

For the mmDNA base pairs modeled in this study, containing Ag^{+}, Hg^{2+}, and Au^{+}, the fully occupied d-orbitals create a stable state without spin dependence, the unoccupied valence orbital

forming an antibonding valence orbital with the nearby bases. Spin-orbit coupling was not considered in these models, which would likely have a non-negligible effect due to the high atomic weight of the metals, especially with nanowire geometries that contain multiple ions. Initial calculations suggest that larger mmDNA chains can create large, delocalized orbitals for electron transmission, which can be investigated with decoherent transport models in a nanodevice.

4. Metal Mediated DNA Force Field Development

The programmable self-assembly characteristics of DNA render it an optimal platform for fabricating nanoscale architectures with precise spatial control and structural versatility.[66] Further customizability has been demonstrated with modification of the nucleotides themselves, ranging from modification of the bases[9] to incorporation of custom functional groups.[67] Such modifications exploit the inherent Watson-Crick hydrogen bonding patterns to generate stable scaffolds. As an alternative, a transition metal can be used to mediate bonding between bases, initially observed between complementary bases[68] as well as between pairs of matching bases.[69–71] The incorporation of metal centers has enabled the development of DNA nanowires with higher measured conductivities than their non-metallated counterparts.[57] The library of metal-mediated DNA (mmDNA) basepairs has been expanded and extensively characterized, with recent work demonstrating that the electronic properties are highly dependent on the metal location and ligand modification.[72]

While the synthesis and characterization of mmDNA has been studied extensively over the past 20 years, knowledge gaps remain in the structural properties of mmDNA heterostructures.[73] X-ray crystallography has been used in previous works to determine the structure of DNA bases, with resolutions ranging from 2-5Å.[64,72,74] These methods generate 3D maps of electron density from the periodic structure, which are fitted to the known DNA base sequence to predict the structure.

However, the crystal structure itself can physically constrain atom movement, leading to strained bonding across the structure. Additionally, the temperatures used for crystallization may not match the environment of a constructed nanostructure or device. To address these gaps, ab initio methods have been applied to mmDNA, with initial studies focusing on the electronic properties of metal-nucleobase interactions on single basepairs[75,76] and more recent studies extending this analysis to look at charge transport in a nanodevice.[77] The dynamics of such structures have been previously studied, ranging from molecular dynamics (MD) modeling treating the coordinated metal as a lone ion[78] or directly parameterizing the metal-base bonds[79] to more complex hybrid quantum mechanics/molecular mechanics (QM/MM) treatment.[80,81] As the range of mmDNA basepairs continues to grow,[72,82] developing adaptable force fields for modeling complex mmDNA heterostructures has become essential, particularly for future innovations in nanodevices and biomaterials.

To address this need, we present an automated ab initio modeling approach that generates bonded terms for MD force fields, enabling efficient scaling to larger heterostructures. Rather than relying on structures generated from X-ray crystallography in a highly constrained environment, a custom force field can be constructed to allow for simulation at ambient conditions. This work builds on the method introduced by Seminario *et al.* to use a projection of the atomic Hessian matrix to generate force constants for bond and angle interactions.[83] To accomplish this, we have built a tool that modifies the existing *pyMSMT* software package[84] to allow parameterization of

mmDNA basepairs. In this work, we have tested this tool on the recently characterized cytosine-thymine mismatch basepairs, with two metallic coordination setups: Ag atoms coordinated in two locations (referred to as C-Ag_2-T, PDB ID 7SDS) and a single coordinated Hg atom (referred to as C-Hg-T, PDB ID 7SD7).[72] These were chosen due to their use of standard Watson-Crick nucleobases that are well characterized and they both contain similar metal binding sites (N3 nitrogen of Cytosine and Thymine). In addition, the C-Ag_2-T base contains metal in the major groove of the DNA, a metal modification that has not been modeled previously. For comparison, we have considered the complementary case with an A-T basepair, similarly characterized by X-ray crystallography (PDB ID 5W6W),[85] and a C-T mismatch case, without metals (referred to as the CT mismatch strand, initial structure based on PDB ID 7SDS).[72] Using the developed force fields allowed us to analyze the mmDNA electronic behavior in ambient conditions to better understand the electronic properties of constructed mmDNA nanodevices, applying the force fields to structures with more than one metal-modified DNA basepair.

a. Modeling Methods

The experimentally obtained X-ray crystallography tensegrity triangles were considered as the starting point for force field generation.[72] A flow chart for the process flow is shown in Figure 4.1, starting out with generation of the custom residue input files for the DFT calculation. First, seven basepair structures were extracted from these large PDB structures for analysis. Then, hydrogens were added based on the modified parmBSC1 force field, following the procedure used in previous

work.[86] These structures were used as the input file to the modified MCPB.py tool,[87] which constructed Gaussian input files for calculations. After running these generated input files in the Gaussian DFT package, the output files were again plugged into the modified MCPB.py tool to generate the force field parameters. The modified MCPB.py tool has been released with this work as a Python package building on previous implementations.[88,89]

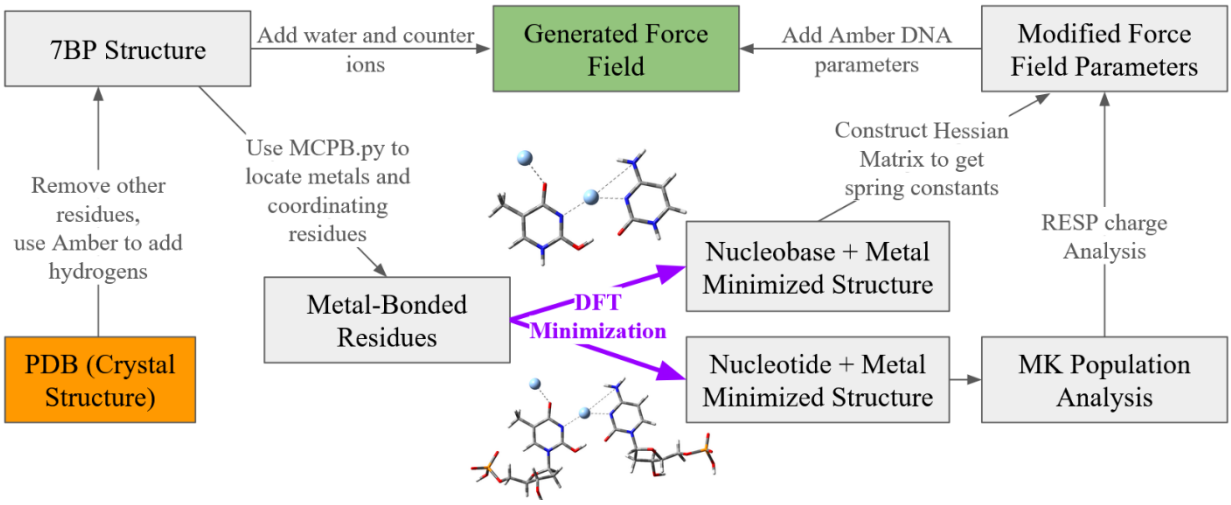


Figure 4.1 - The flow chart for generation of the force fields, starting with the X-ray crystallography reference structure and using the patched MCPB.py to generate parameters.

Ab initio DFT calculations were conducted using the Gaussian 16 software package to minimize the metallated basepair, calculate the Hessian matrix, and determine atomic charges with the MK population method.[90] The B3LYP functional was used for all calculations with a mixed basis set, using the 6-31G** Pope basis set[91] for non-metal atoms and the LANL2DZ basis set with effective core potentials[92] for metal atoms. A polarizable continuum model with a dielectric constant of 78, which is representative of water, was used in the DFT calculations to find the atomic charge and

generate the force field parameters.[93] After DFT optimization on the metallated nucleobase and nucleotide structures, the optimized atomic coordinates were used for the frequency calculation and atomic charge calculation respectively.

In this study, the bonded parameters were derived from DFT calculations using the proposed modified MCPB.py workflow. Specifically, the values for k_b, k_θ, r_0 and θ_0 shown in equation 4 (see chapter II) were parameterized between the coordinated metals and DNA atoms (e.g., N3 of cytosine and thymine), which were incorporated into the modified parmBSC1 force field. Dihedral parameters were not parameterized and set to zero for ligand-metal bonds. The Coulomb charges in equation 5 (see chapter II), q_i and q_j, were parameterized by calculating charges on the metal atom and coordinating residues. The LJ parameters were assigned from previously published values fitted to the OPC water model, using parameters for the monovalent Ag^+ ion, which are $R_{min} = 1.316$ Å and $\epsilon = 0.00602547$ kcal/mol,[94] and for the divalent Hg^{2+} ion, which are $R_{min} = 1.305$ Å and $\epsilon = 0.00523385$ kcal/mol.[95] By focusing on the bonded terms for key metal-DNA interactions and leveraging non-bonded terms for long range effects, the generated force field achieves accurate modeling of structural stability of metallated DNA systems and, by extension, their electronic behavior.

Equilibrium bond lengths were extracted from the nucleobase DFT minimization and force constants were derived from the generated Hessian matrix, which is given by the tensor

$$H_{i,j} = \begin{bmatrix} \frac{\partial^2 E}{\partial x_i \partial x_j} & \frac{\partial^2 E}{\partial x_i \partial y_j} & \frac{\partial^2 E}{\partial x_i \partial z_j} \\ \frac{\partial^2 E}{\partial y_i \partial x_j} & \frac{\partial^2 E}{\partial y_i \partial y_j} & \frac{\partial^2 E}{\partial z_i \partial y_j} \\ \frac{\partial^2 E}{\partial z_i \partial x_j} & \frac{\partial^2 E}{\partial z_i \partial y_j} & \frac{\partial^2 E}{\partial z_i \partial z_j} \end{bmatrix} \quad (38)$$

in cartesian coordinates for atoms i and j, which are projected onto the internal coordinate system specified by bonds and angles. The atomic charges used in the Coulomb term of equation 2 were calculated using the Merz-Singh-Kollman scheme to calculate atomic charges and fitting them using the restrained electrostatic potential (RESP) approach.[96] As a final step, the MCPB.py program generated a tleap input file that was used to generate the parameter and topology files used for AMBER simulations.

To compare the performance of the generated force field, the initial structures were optimized with a partial DFT energy minimization procedure used in a Chapter 3.[72] This method is referred to in this work as the *partial minimization* method and exclusively uses DFT calculations to generate a structure. In this method, a force calculation was conducted to locate atoms that exceeded the specified force threshold, 5 eV/Å. Then, the energy minimization DFT calculation was applied to these unrestrained atoms, keeping the rest of the structure frozen. The final structure was subjected to a force calculation applied to all atoms, and the cycle was repeated until all atoms were below the 5 eV/Å cutoff.

This study built on the current version of the MCPB.py toolbox,[97] which contains force field parameters for proteins as well as the GAFF parameters for small organic molecules. The existing MCPB.py toolbox was modified to include parameters for the DNA, a key contribution of this work. The parmBSC1 force field was added to the toolbox and updated to include the modified thymine residue (residue label DTM) that coordinates with the metal. A comprehensive mol2 file containing the structure of each DNA residue was generated to be read by MCPB.py as a library of atom types within each residue. Using the modified force field files, the MCPB.py tool was able to locate which specific residues contain metal-bonding atoms (within the specified cutoff distance) to generate DFT input files.

Two different sized structures, the metallated nucleobase and nucleotide structures, were used as input files for Gaussian calculations. In the updated software package, the nucleobase structure is generated without the backbone and is labeled as the "small" structure, to be used for unconstrained DFT minimization and Hessian matrix generation. The nucleotide structure containing the full backbone is used for the "large" case, which was used for DFT minimization constraining the phosphate atoms before charges were read using the Merz-Kollman-Besler population method.[90] For this "large" case, the 5' and 3' ends of the residue are capped with hydrogens and the system charge was calculated based on the number of phosphates and oxidation state of the metal atoms.. A full list of files as well as changelogs with specific changes made to the pyMSMT package are available in the GitHub Github repository.[87] A summary of the modifications made to each file are given below:

File	**Modification**
lib.py	• Added the modified parmBSC1 force field information
element.py	• Added the modified parmBSC1 force field information
gauio.py	• Modified the atom counting method to allow for hydrogen/ligand capping of the residues before • Changed default basis/options for gaussian input files
amber_modeling.py	• Added backbone connection to tleap input file for modified bases
gene_model_files.py	• Added carbon bond option • Updated residue counter to not conflict with DX3 and DX5 residue names • Added methods for constructing base and base+backbone fragment structures for DFT analysis (both with proper hydrogen/oxygen termination)
mol.py	• Created basepair terminus classification for residue object
pdbio.py	• Added “TER” keywords for DNA strands
MCPB.py	• Added “include_carbon” option for force field generation

The AMBER 20 software package was used to run MD simulations using each of the generated force fields for the seven basepair structures. The MD simulations were conducted in five stages. First, the system underwent a constrained minimization, relaxing the solvent and ions while constraining the DNA and metal atoms (restraining force constant set to 50 kcal mol^{-1} $Å^{-2}$). Then, a second minimization stage was applied to the entire system removing restraints. Both

minimizations used a non-bonded cutoff of 10.0 Å and ran for 5000 cycles each. The third heating stage raised the system's temperature from 0 K to 300 K over 1 ns under constant volume (NVT) conditions, with restraints applied to the DNA and metal atom's structure. Following heating, the system was further relaxed in a fourth equilibration stage with gradually reduced restraints conducted under constant pressure (NPT). The fifth (and final) stage was a production run, conducted under constant pressure (NPT) conditions without restraints for 100 ns. The SHAKE algorithm[98] was applied throughout to constrain bonds involving hydrogen atoms, and the Particle Mesh Ewald (PME) method[99] was used for long-range electrostatics. To analyze the structural dynamics of the DNA and metal atoms structures, we measured the RMSD using VMD[100] to observe the conformation changes during the MD simulation.

MD minimization was also used to conduct the "restrained minimization" cases applied to the single C-Ag_2-T and C-Hg-T structures. For this case, the modified parmBSC1 force field[101] was used for the DNA bases while fitted non-bonding parameters were used for Ag and Hg ions.[102,103] Instead of adding bonding parameters to parameterize the distances, harmonic restraints were applied to the N3-AG/HG-N3 bond angle and the O4-AG bond distances. These were only applied during the second stage of minimization, using a coefficient of 400 kcal/mol per square unit (°/Å) as an energy penalty during 5000 cycles of minimization.

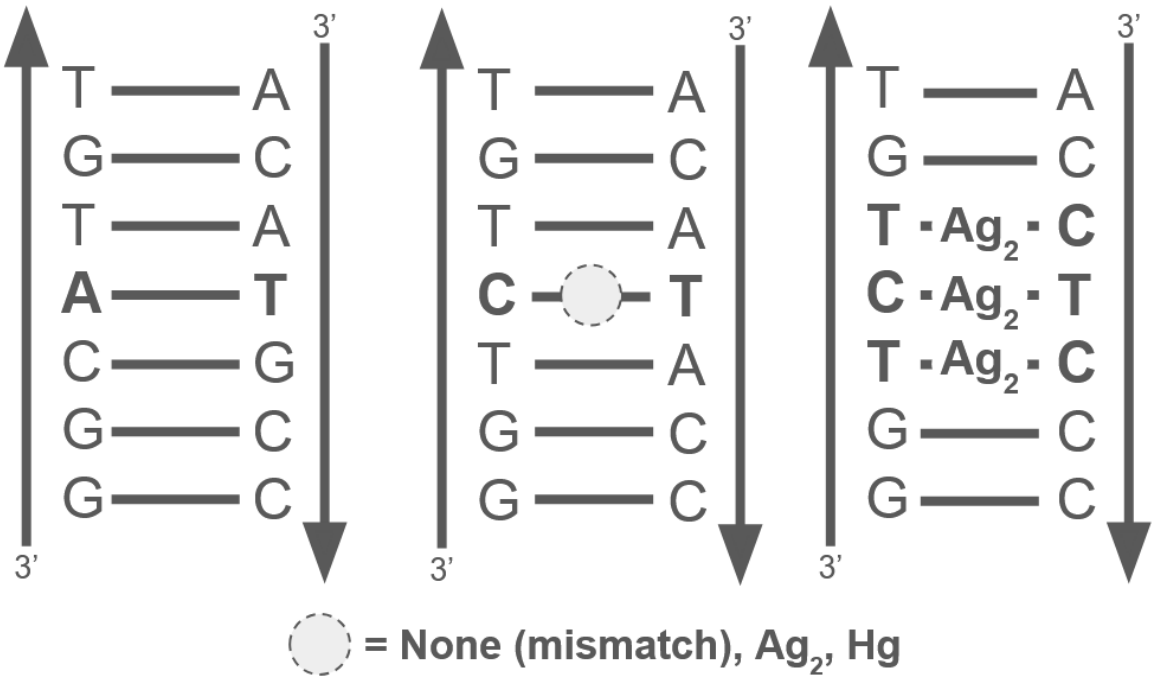


Figure 4.2 - Sequence information for each of the five different cases. The leftmost sequence is the complementary case, with an AT central basepair. The middle sequence was used for the three cases: single C-T mismatch, single C-Ag_2-T and single C-Hg-T. The right sequence was used for the triple C-Ag_2-T case.

b. Results

Force fields were constructed for the 7 basepair C-Ag_2-T and C-Hg-T sequences, as depicted in Figure 4.2. To probe the effects of interactions between multiple mmDNA basepairs, a triple C-Ag_2-T case was constructed for parameterization by modifying the basepairs 3 to 5 in each sequence, with the coordinated pair of Ag atoms placed at each of the three C-T mismatches. All structures were then subjected to energy minimization followed by 100ns MD simulation at room

temperature, analyzing structural stability and conformation changes across the simulated trajectories.

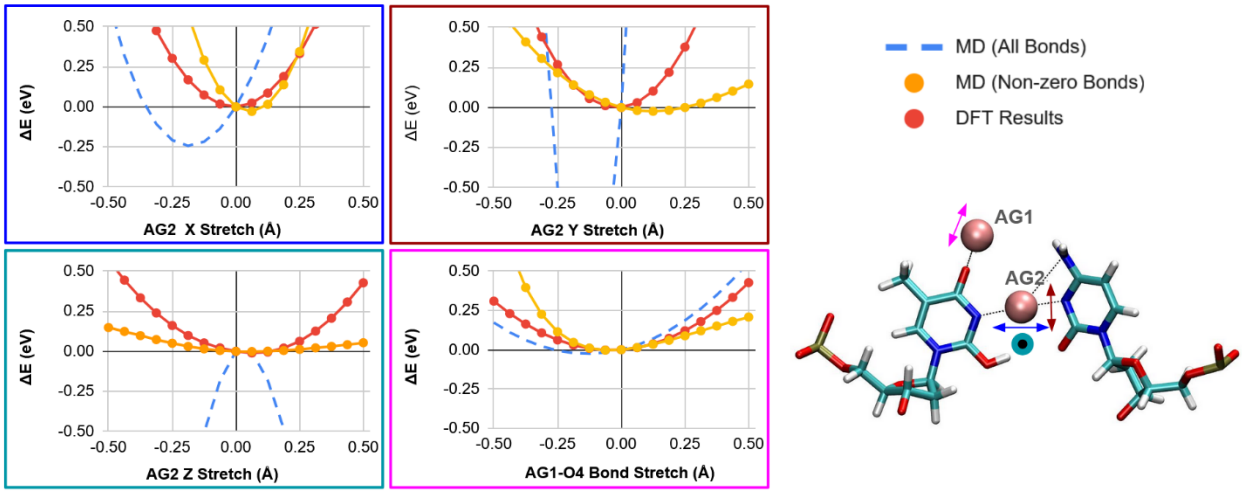


Figure 4.3 - Energy sampling results for the C-Ag$_2$-T single basepair, with color-coded graphs on the left matching up with Ag atom movement given by the colored arrow on the right. Three energy profiles are shown for each case: MD energy with all non-H bonds parameterized (MD All bonds), MD energy with just non-zero N and O bonding parameters (MD Non-zero bonds), and DFT energy using a solvent model. AG2 refers to the coordinated Ag atom and AG1 refers to the major groove AG atom. The equilibrium bond lengths and bond angles are in following section.

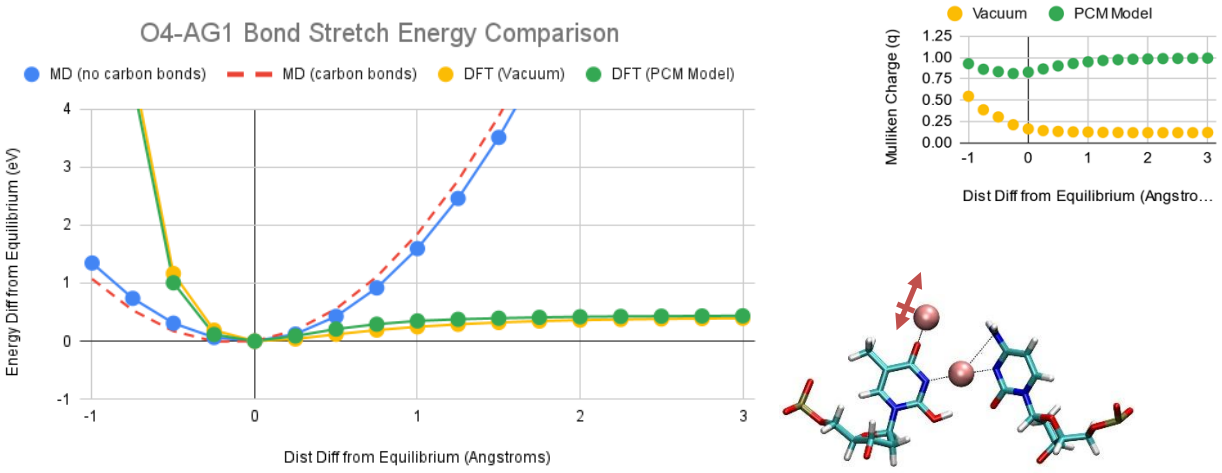
Force Field Generation

Force field parameters were derived from the Hessian matrix calculated after DFT minimization of the mmDNA sites. Initially, all non-hydrogen atoms (C, N, and O) that were within the set bond length cutoff (3.0Å) were used to generate bond and angle force constants in an attempt to use all the information from the Hessian matrix for force field generation. To better understand the limits of this fully parameterized model, the computed force field was compared directly to ab initio

modeling to sample the potential energy profile of the metal-base interactions. A comparison is shown on a shorter length scale in Figure 4.3, sampling energies from moving the coordinated Ag in 3 directions and the major groove Ag atom along the Ag-O4 bond length. It was found that including bonding parameters for all non-hydrogen atoms within the cutoff distance led to an overconstrained system that strayed significantly from the DFT sampling results, whereas only including non-zero nitrogen and oxygen bond force constants matched up well with the DFT results. Not only did adding bonds to the force field overconstrain the interactions, but also it led to scaling or removal of important non-bonded interactions for neighboring atoms, based on the implementation of the AMBER force field.[104] In the final parameterization, only nitrogen and oxygen bonds were included, removing bonds that were assigned a force constant of zero. All these parameters calculated for the single C-Ag2-T basepair and single C-Hg-T basepair are listed in the following section.

In the generated force field, the most significant metal-base interaction was between the coordinated metal (Ag or Hg) and the N3 nitrogen from both the cytosine and thymine, which coordinated the metal with symmetric bond force constants in the range of 25-35 kcal/(mol•Å) for Hg and 50-75 kcal/(mol•Å) for Ag. This range is similar to bond force constants calculated for other systems with closed-shell metal interactions, such as Cu binding enzymes (40-45 kcal/(mol•Å) for Cu-S bonds)[105] or those used in the generalized zinc AMBER force field (30-120 kcal/(mol•Å) for Zn-N bonds).[88] In the Ag basepair, the major groove Ag atom bonded to only the thymine O4 atom with a significant bonding force constant and with negligible or low angle force

constant (see the following section). These bond force constants remained the same even when the metal atoms were modeled together in the three metal basepair stack. Initial modeling of the triple basepair stack used a cutoff of 3.4 Angstroms, equal to the average spacing between basepairs, to calculate force constants for inter-basepair Ag-Ag bonding. Since the AMBER force field does not include the non-bonded terms for nearest neighbors or second nearest neighbor terms,[104] removing these bonds ensured that the inter-basepair interactions used the non-bonded Lennard Jones and Coulomb force field terms from fitted charges instead.



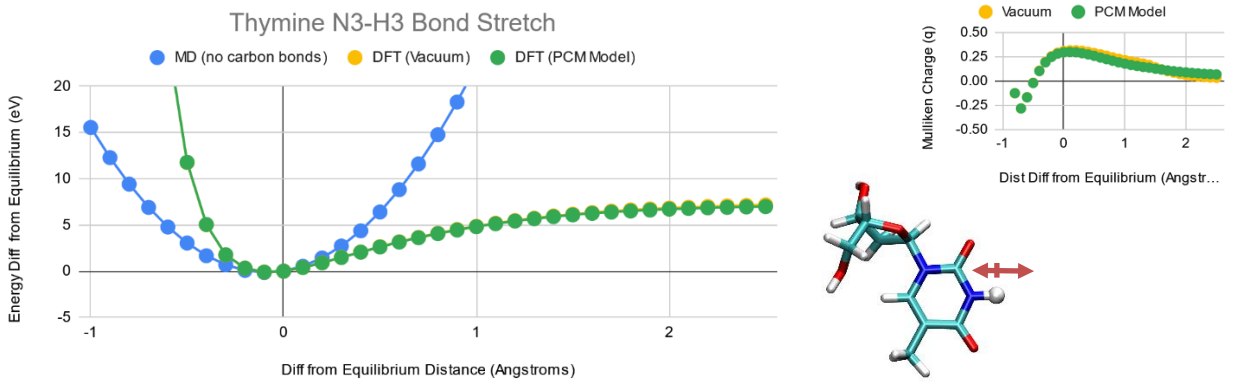


Figure 4.4 - Energy sampling comparison between the generated thymine O4-AG1 bond parameters (above) and the parmBSC1 thymine N3-H3 (below) bond parameters. Mulliken charges of AG1 and H3 atoms included as inset graphs on the left and right, respectively.

While the metal-base bonding coefficient values are much smaller than those for nearby organic covalent interactions (see the italicized entries in the force field parameter tables in the following section), the bonding parameters stabilize the metal atoms enough to generate stable conformations. The limits of the bonded vs non-bonded interactions are quite clear when looking at larger bond distances. Sampling these bond distances, there is significant deviation between the DFT and MD energies. Specifically, the energetics of major groove Ag atom in C-Ag_2-T and the N3-bonded hydrogen atom in thymine highlights the limitations of the second order (parabolic) approximation of MD at larger length scales. This difference is illustrated in Figure 4.4 with the DFT energy flattening at larger length scale compared to the quadratic relationship from the AMBER force field. This shows the limits of the second order approximation for bond descriptions, but a closer analysis also highlights the limits to the DFT modeling at these longer length scales. For instance, the net charge on the thymine H3 atom across the trajectory does not

approach the expected value of +1 even when the atom is moved 2 Angstroms away from the equilibrium position. This demonstrates a breakdown of the mean-field approximation and spin model, with the DFT method forcing a closed-shell orbital to extend between atoms that should form an ionic or radical pair. This was also found to be true for the Ag atom when modeled with the implicit water solvent model, although modeling in a water solvent did generate a +1 Mulliken charge after 1 Angstrom of stretching.

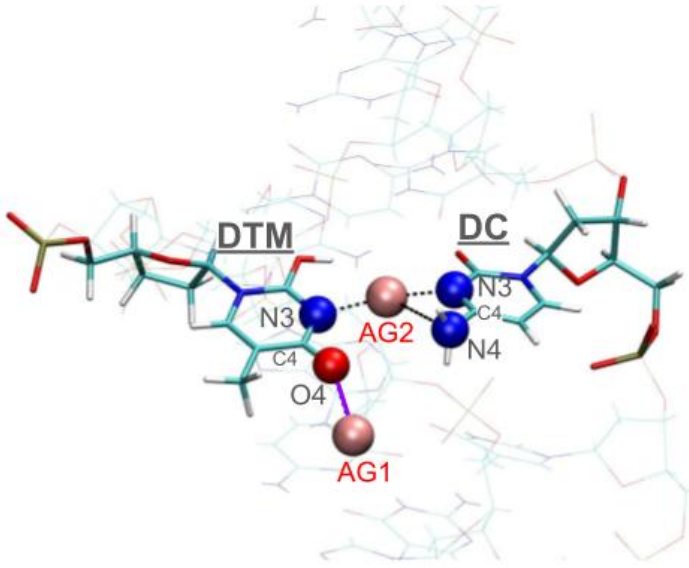
Table 1 - A comparison of DFT energies derived from the seven basepair structures, starting with the original X-ray structure and including the 3 minimization methods applied.

System	Property	X-ray Structure	Partial DFT Min	Angle Restrained Minimization	**Generated Force Field**
C-Ag_2-T	HOMO/LUMO Gap	3.44 eV	3.48 eV	3.55 eV	**3.31 eV**
	HOMO Energy	-5.20 eV	-5.26 eV	-5.18 eV	**-5.21 eV**
	Relative DFT Energy	0.00 eV	-20.06 eV	-21.35 eV	**-23.86 eV**
C-Hg-T	HOMO/LUMO Gap	3.39 eV	3.39 eV	3.55 eV	**2.79 eV**
	HOMO Energy (eV)	-5.15 eV	-5.15 eV	-5.18 eV	**-5.25 eV**
	Relative DFT Energy	0.00 eV	-0.29 eV	-3.60 eV	**-7.54 eV**

Generated Force Field Parameters

Single C-Ag$_2$-T

Metal Atom	Partial Charge (q)
AG1	+0.91
AG2	+0.51



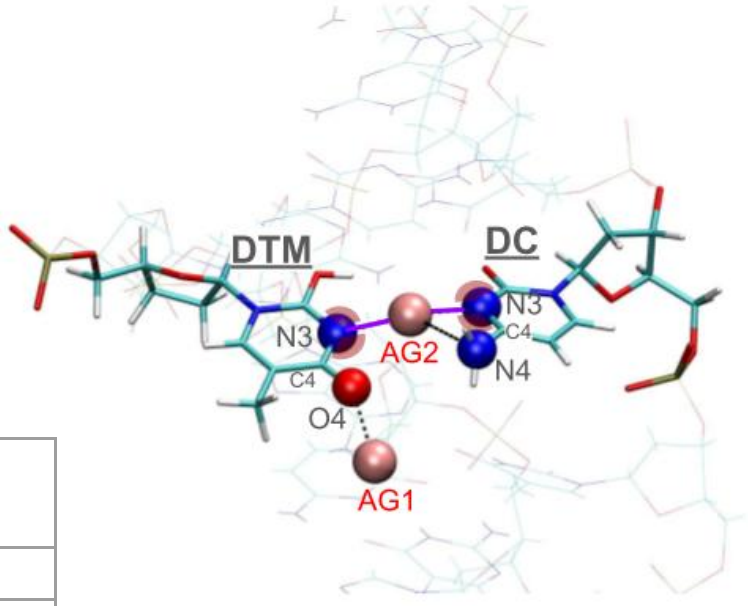
Major Groove AG (AG1)

Bond	Equilibrium Distance (Å)	**Spring Constant (kcal mol^{-1}Å^{-2})**
C5-C7	*1.51*	***317.0***
O4-AG1	2.28	**35.0**

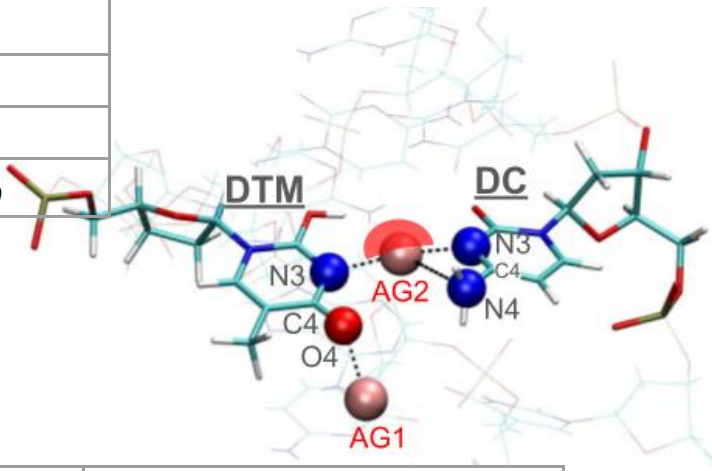
Bond	Equilibrium Angle (°)	**Spring Constant (kcal mol^{-1} rad^{-2})**
C4-O4-AG1	140.1	**19.7**
C2-O2-H	*109.5*	***35.0***

Coordinated AG (AG1)

Bond	Equilibrium Distance (Å)	**Spring Constant (kcal mol^{-1}Å^{-2})**
N4-C4	*1.34*	***481.0***
N4-AG2	3.42	**3.1**
(DTM) N3-AG2	2.20	**65.1**
(DC) N3-AG2	2.23	**53.7**



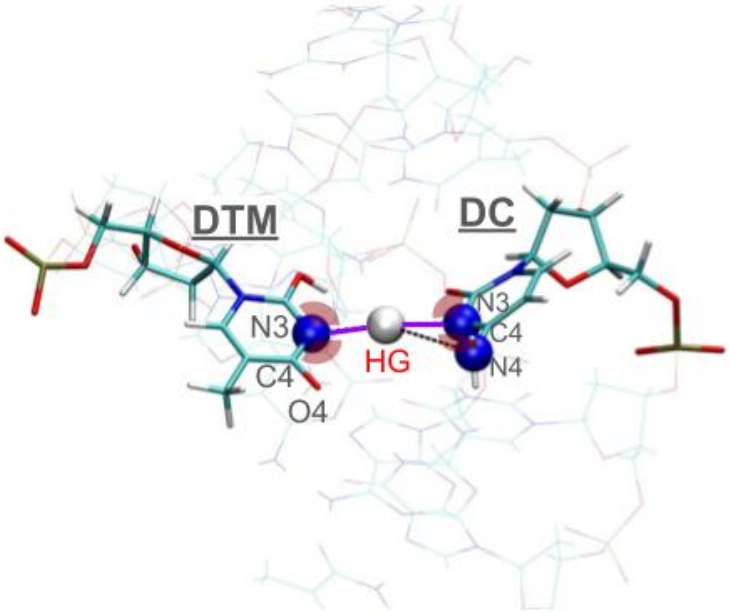
Bond	Equilibrium Angle (°)	**Spring Constant (kcal mol^{-1} rad^{-2})**
(DC) H-N4-AG2	109.7	**25.8**
(DC) C4-N4-AG2	70.2	**31.7**
(DTM) C2-N3-AG2 **(DTM) C4-N3-AG2**	119.8	**89.5**
(DC) C2-N3-AG2	109.3	**95.3**
(DC) C4-N3-AG2	129.3	**92.3**
N3-C4-O4	*122.5*	***80.0***



Bond	Equilibrium Angle (°)	**Spring Constant (kcal mol^{-1} rad^{-2})**
(DC) N3-AG2-N4 (DC)	41.8	**29.4**
(DTM) N3-AG2-N4 (DC)	136.1	**29.3**
(DTM) N3-AG2-N3 (DC)	177.2	**44.7**
N3-C4-O4	*122.5*	***80.0***

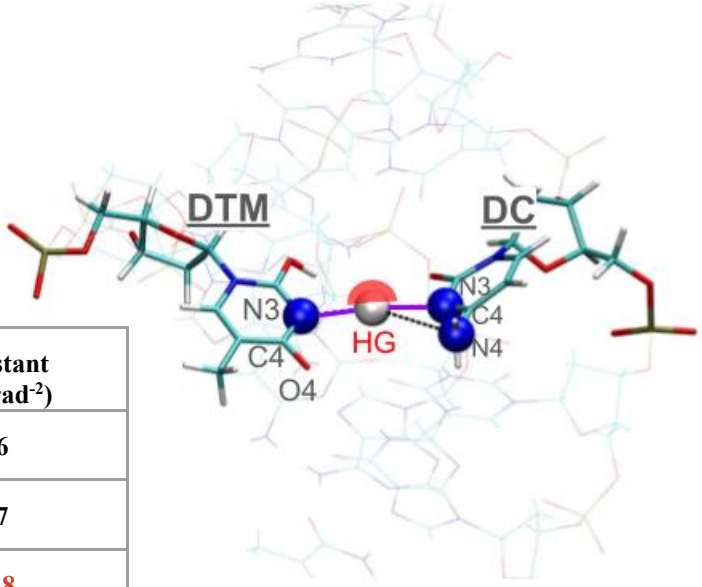
Single C-Hg-T

Metal Atom	Partial Charge (q)
HG	+1.49



Bond	Equilibrium Distance (Å)	**Spring Constant (kcal mol^{-1}Å^{-2})**
N4-C4	*1.34*	***481.0***
N4-HG	3.64	**3.7**
(DTM) N3-HG	2.51	**25.3**
(DC) N3-HG	2.39	**35.4**

Bond	Equilibrium Angle (°)	**Spring Constant (kcal mol^{-1} rad^{-2})**
H-N4-HG	109.7	**30.5**
C4-N4-HG	70.7	**38.7**
(DC) C4-N3-HG	131.7	**87.0**
(DTM) C4-N3-HG **(DTM) C2-N3-HG**	119.7	**113.7**
(DC) C2-N3-HG	106.5	**90.0**
N3-C4-O4	*122.5*	***80.0***



Bond	Equilibrium Angle (°)	**Spring Constant (kcal mol^{-1} rad^{-2})**
(DTM) N3-HG-N4 (DC)	145.55	**35.6**
(DC) N4-HG-N3 (DC)	38.5	**35.7**
(DTM) N3-HG-N3 (DC)	175.9	**103.8**
N3-C4-O4	*122.5*	***80.0***

DFT Energy Analysis

Applying these force fields to the full seven basepair strand geometries (see Figure 2), the trend in DFT energies shown in Table 1 indicates that the force fields work as expected for full DNA strands. The results from these four cases (crystal structure, partial DFT minimization, restrained minimization, and the force field generated in this work) are shown in Table 1. In both the Ag and Hg case, the structure determined by the generated force field gave the lowest DFT energy. It should be noted that the restrained minimization method was quite effective in both cases, reducing the DFT energy significantly, but the structures were found to be highly unstable during the equilibration and production MD runs and exploded after a few nanoseconds (data not shown).

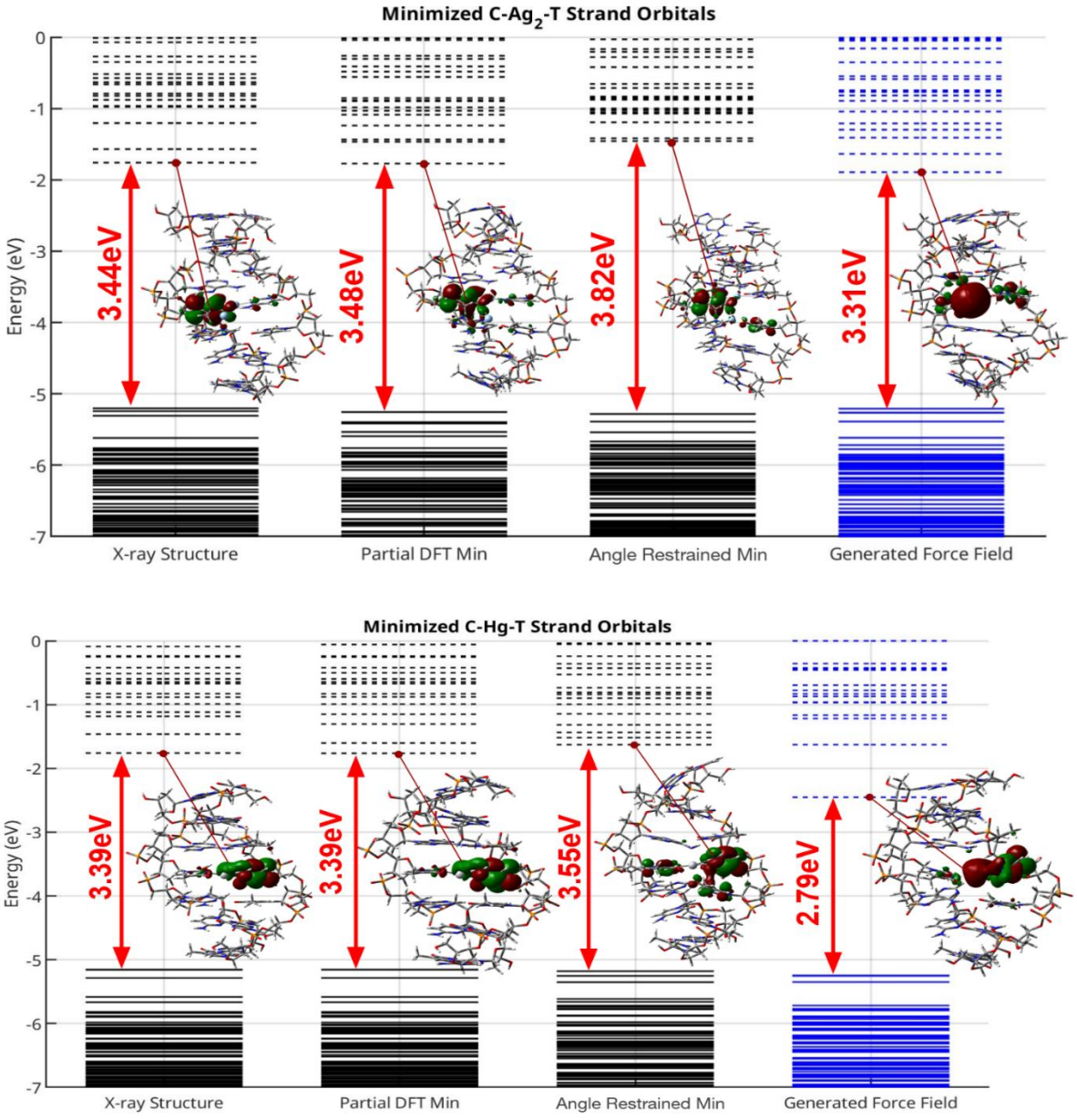


Figure 4.5 - Side by side comparison of the seven basepair structures for the C-Ag_2-T (top) and C-Hg-T (bottom) basepairs before and after minimization. The HOMO-LUMO gap is labeled in red for each band structure with the force field minimized case highlighted in blue.

The band gaps of the structures remained relatively constant after minimization, but the LUMO orbitals were found to shift significantly after minimization with the generated force field. In the C-Ag_2-T case, the LUMO orbital was delocalized over the major groove Ag atom after minimization. In the C-Ag_2-T case, the LUMO orbital was delocalized over the major groove Ag atom after minimization. For the C-Hg-T case, the LUMO orbital showed more occupation of the coordinated Hg atom in addition to the surrounding basepairs. Furthermore, in the C-Hg-T case the HOMO-LUMO gap dropped significantly (from 3.39 to 2.79 eV) after minimization using the generated force field. The LUMO orbitals and corresponding band structures are shown side by side in Figure 4.5, with the force field minimized case highlighted in blue.

Expanding the structure to the triple C-Ag_2-T basepair case, three approaches were applied to parameterization, with force field minimization results compared to the structure before minimization. These approaches, which varied based on parameterization constraints, are as follows: (i) parameters from the single C-Ag_2-T structure applied to all three C-Ag_2-T basepairs, (ii) parameters gathered from minimizing the three C-Ag_2-T basepairs together including basepair-basepair bonding, and (iii) parameters gathered from minimizing the three C-Ag_2-T basepairs together excluding basepair-basepair bonding. Applying the two-step minimization in each of these cases, the parameters in case (i) led to the highest energy drop, a total of 15.84 eV energy difference after minimization, and only these parameters are given in the previous section. The final structures for each of these cases are depicted in Figure 4.6, showing that parameterization

of case (ii) created significantly more strain between neighboring basepairs. While the bonding parameters generated from case (iii) matched up closely with those generated for the single C-Ag_2-T basepair, the N3-Ag-N3 angle changed significantly (dropping from 177 degrees to 140 degrees), leading to distortion of the basepair after minimization. The HOMO-LUMO gap increased significantly after minimization from 1.92 eV to 2.77 eV for case (i). Orbital energies and LUMO orbitals for each of the structures are plotted below in Figure 25.

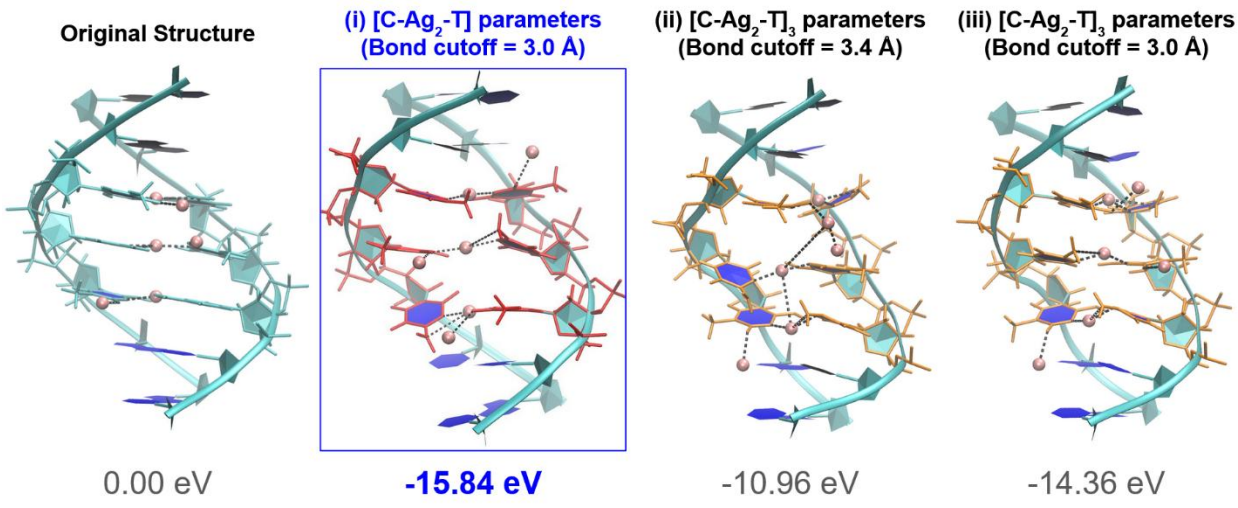


Figure 4.6 - The constructed triple C-Ag2-T basepair structure before and after minimization, comparing the three parameterization cases (i)-(iii), showing all new metal bonds parameterized with dotted lines. Total DFT energy relative to the original structure is labeled below each structure, The chosen force field set is case (i), labeled in blue, with the lowest DFT energy after minimization.

Three parameterization methods were tested for the Triple C-Ag_2-T basepair case. Carbon bonds were not parameterized in any of the cases. The cases are as follows, referenced in the orbital energy diagram in Figure 4.6 and Figure 4.7:

Case i. MCPB.py used to generate force field using the Hessian matrix and minimized structure for the for a single C-Ag_2-T base with bonding cutoff set to 3.0 Å, only bonds with non-zero bonding constant parameterized.

Case ii. MCPB.py used to generate force field using the Hessian matrix and minimized structure for the three alternating C-Ag_2-T basepairs with bonding cutoff set to 3.4 Å, all possible bonds between Ag, N, and O parameterized.

Case iii. MCPB.py used to generate force field using the Hessian matrix and minimized structure for the three alternating C-Ag_2-T basepairs with bonding cutoff set to 3.0 Å, only Ag bonds to N and O with non-zero bonding constant parameterized.

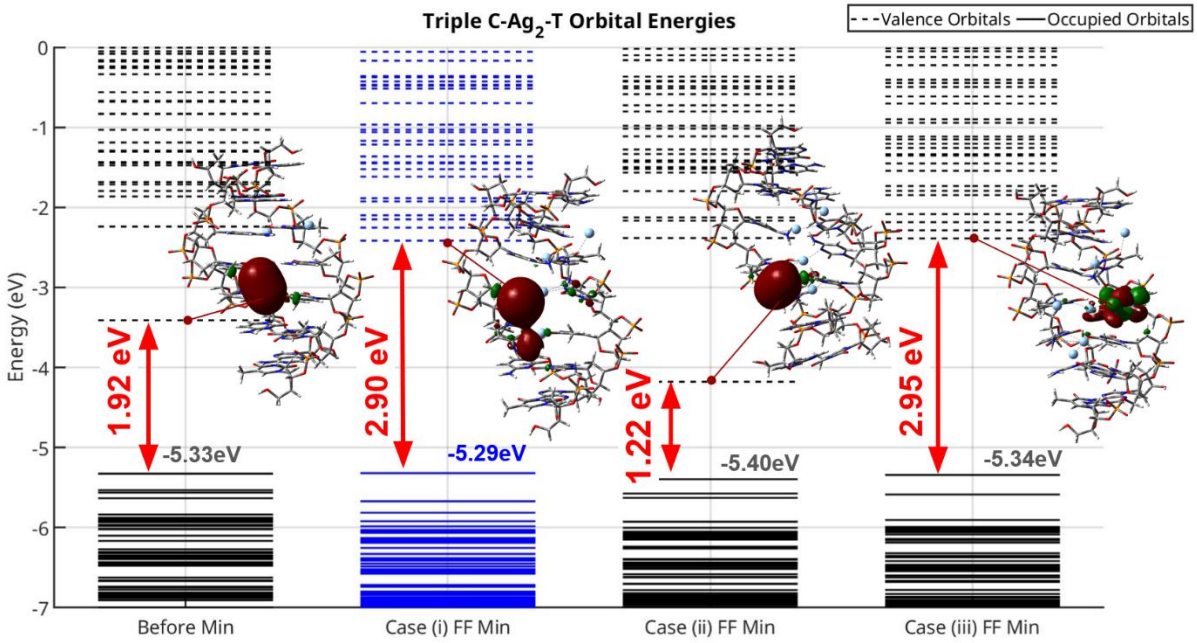


Figure 4.7 - Orbital energies diagram for the triple C-Ag$_2$-T structure before and after minimization for all 3 parameterization methods (case i-iii). HOMO-LUMO gaps are labeled for each case in red, with an isosurface of the LUMO orbital included as an inset to each plot ($|\Psi|$=0.02 plotted, negative wave function is red, positive wave function is green)

Structure Dynamics

The constructed force fields were used to generate molecular dynamics trajectories starting with a two-step minimization run followed by equilibration run and then a 100 ns simulation. In each case the structure remained stable, with the root mean square deviation (RMSD) calculated only over the middle three basepairs of the atomic positions in the metal modified strands and the mismatch case matching up with the standard complementary strand. The RMSD between frames is visualized in Figure 4.8 showing the 3 metal modified cases, the single C-T mismatch case, and the complementary case (A-T central basepair). The complementary case has RMSD values mostly

between 1.5 and 2 Å with some regions showing lower values that suggest partial stabilization. The single C-T mismatch system has higher fluctuations, which indicates reduced stability compared to the complementary case with A-T basepair. The single C-Hg-T case has RMSD values between 1.0 and 2.0 Å, showing improved structural stability compared to the complementary and C-T mismatch cases. The single C-Ag_2-T system is relatively stable, with RMSD values mostly between 1.0 and 2.0 Å. The triple C-Ag_2-T system has the lowest fluctuations, with RMSD values between 0.0 and 1.0 Å during most of the MD simulation. This indicates a significant improvement in structural stability, confirming that adding metal atoms improves DNA stability. The triple C-Ag_2-T system demonstrates that increasing the number of metal atoms further enhances structural stability and reduces fluctuations in mismatched basepairs.

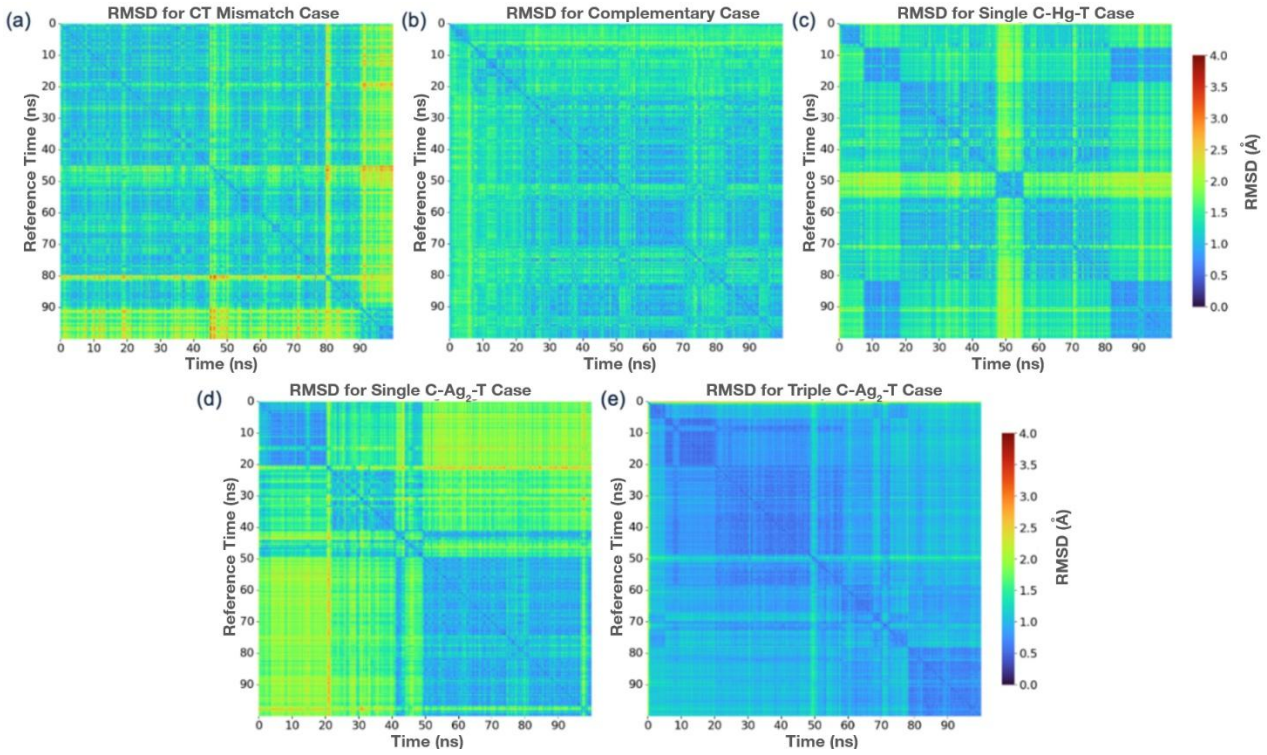
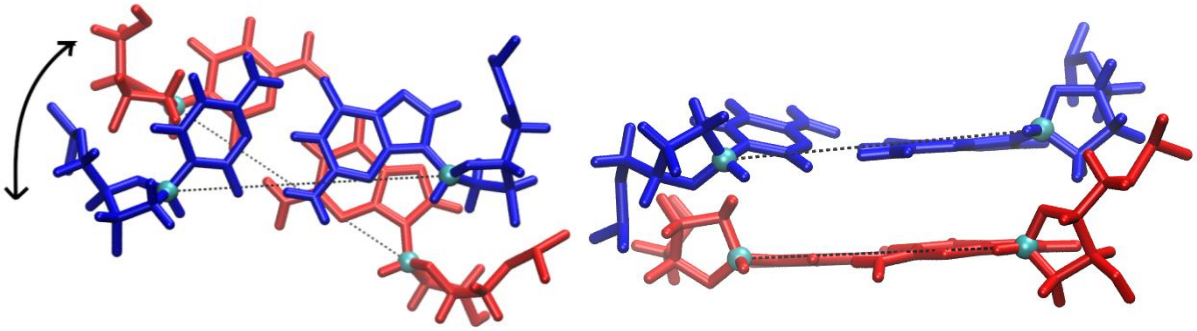


Figure 4.8 - 2D plot of RMSD between simulation frames for the middle three basepairs of the DNA strands across the 100 ns MD simulation for all cases: (a) single CT mismatch, (b) complementary, (c) single C-Hg-T (d) single C-Ag_2-T basepair and (e) triple C-Ag_2-T

To understand the overall DNA structure across the trajectories, inter-basepair angles and distances were calculated across the 100 ns run by saving the coordinates of 100,000 frames that were equally spaced in time. The twist angle between neighboring basepairs were calculated and the results for the central five basepairs are shown in Figure 4.10a. The twist angle for the non-metallated case matched more closely with that of B-DNA (~34°) while the metallated cases had much more variation across the strands. This was found to be especially true for the triple C-Ag_2-

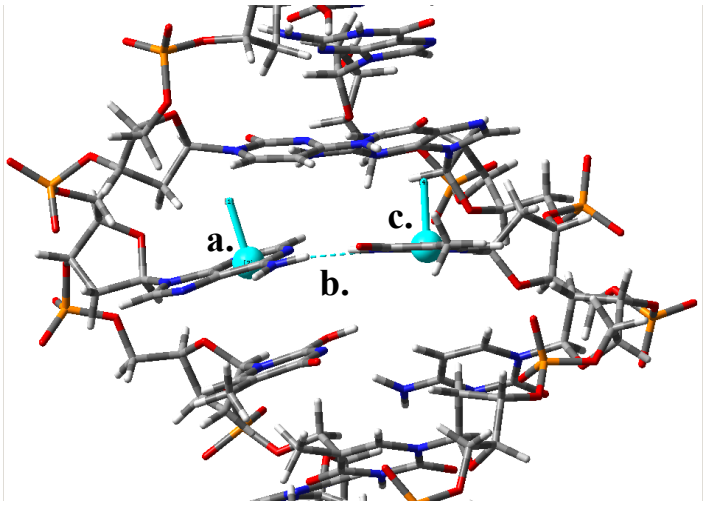
T case, which had a higher average twist angle between the three central metallated basepairs. The basepair propeller angle was measured by taking the dihedral angle between bases in a basepair, matching up with the propeller angle parameter used in the 3DNA software package.[106] This angle showed similar trends, increasing dramatically for the triple C-Ag_2-T case. The ranges of each of these angles for the central five basepairs across the MD trajectory are plotted in Figure 4.10b with an in-depth description of the calculation methodology presented in the following subsections.

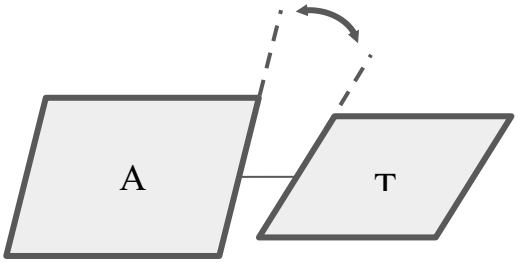
Twist Angle



Twist angle refers to the helical twist angle between neighboring basepairs of DNA (not to be confused with the twist angle between opposite bases). To calculate the twist angle, a reference vector was constructed between C1' atoms in opposing bases. The angle between the reference vector and the next nearest neighbor was reported. Edge basepairs were not included in these measurements due to their instability leading to inaccurate measurements of helical twist angle. Therefore, for the seven basepair structures used in this study, there are four twist angle measurements possible.

Basepair Propeller Angle

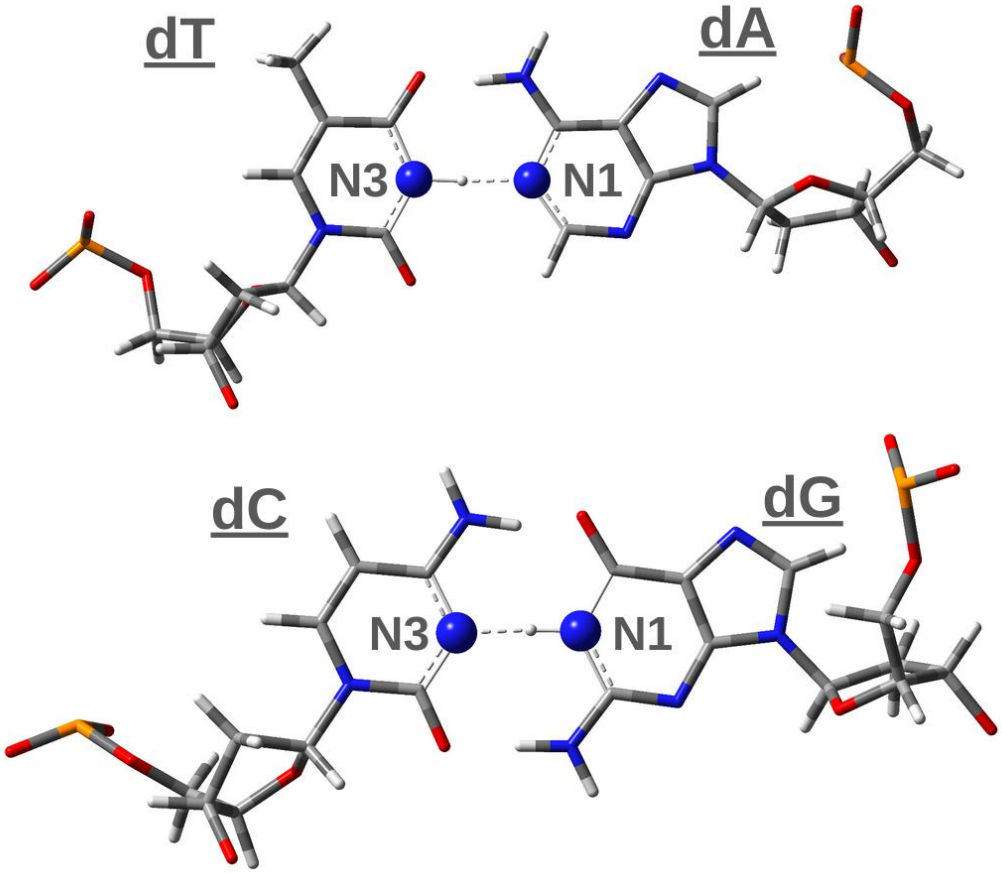




Basepair propeller angles were measured between opposing bases, using a dihedral angle calculated from the normal vector of each nucleobase. First, the normal vector of a characteristic plane was calculated using singular value decomposition. Then, the coordinates of all non-hydrogen atoms in the nucleobase (backbone/sugar excluded) were centered and three principal components were calculated based on their coordinates. The first two principal components span the characteristic plane of the base, with a normal vector given by the third component. These normal vectors were used for vector (a.) and (c.) above, with vector b given by the difference in center of masses for each nucleobase. This vector was calculated by subtracting the weighted average of all non-hydrogen atoms in each nucleobase (backbone/sugar excluded). The dihedral angle was calculated using the atan2 function, allowing the angle to span from -180° to +180°.

Basepair Spacing

Basepair spacing was calculated by measuring the distance between opposing bases in each basepair. Looking at the stability of metal coordination, the N-N distance was calculated based on the following nitrogen locations in each basepair:



This means that for complementary pairs, the distance was measured between the N3 nitrogen on the pyridine (C, T) and the N1 nitrogen on the opposing purine (G, A). Therefore, for the C-T

mismatch (with and without metals) the distance between the opposing N3 nitrogen atoms was calculated.

In addition to the N-N distance, two distances were measured for each basepair: the distance between C1' atoms, the vector used in the helicity calculation above, and the distance between nucleobase centers of mass, the vector used in the basepair propeller angle calculation above. The range in values across the MD trajectory are plotted for the central five basepairs (excluding edge basepairs) for each of the cases in Figure 4.9 below.

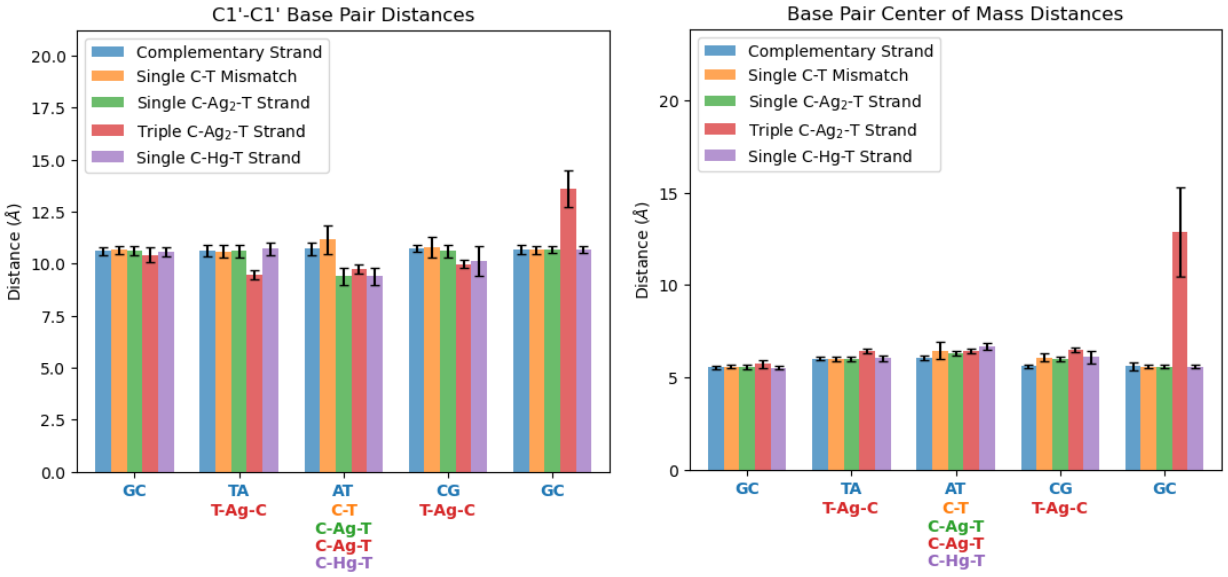


Figure 4.9 - Basepair spacing values for the complementary, single mismatch, single C-Ag2-T, triple C-Ag2-T, and single C-Hg-T cases. Mean is plotted, with error bars given by standard deviation across 100ns trajectory. Labels on X-axis are color coded to match up with non-complementary bases for each of the non-standard DNA cases

Basepair distances were analyzed across the trajectory, specifically monitoring nitrogen coordination of the metal in the mmDNA cases. The distances between the nitrogen atoms of each of the three central basepairs are plotted, with overall ranges and time-dependent curves shown in Figure 4.10c and Figure 4.10d respectively. The metal modified basepairs increased in the N-N distance by 1.4-1.6 Å and showed high stability during MD runs. Only the CT mismatch case shows any instability, with N-N distance varying significantly over the whole 100 ns trajectory. This trend was also observed for the center of mass distances and backbone distances, plotted in Figure 4.9 above.

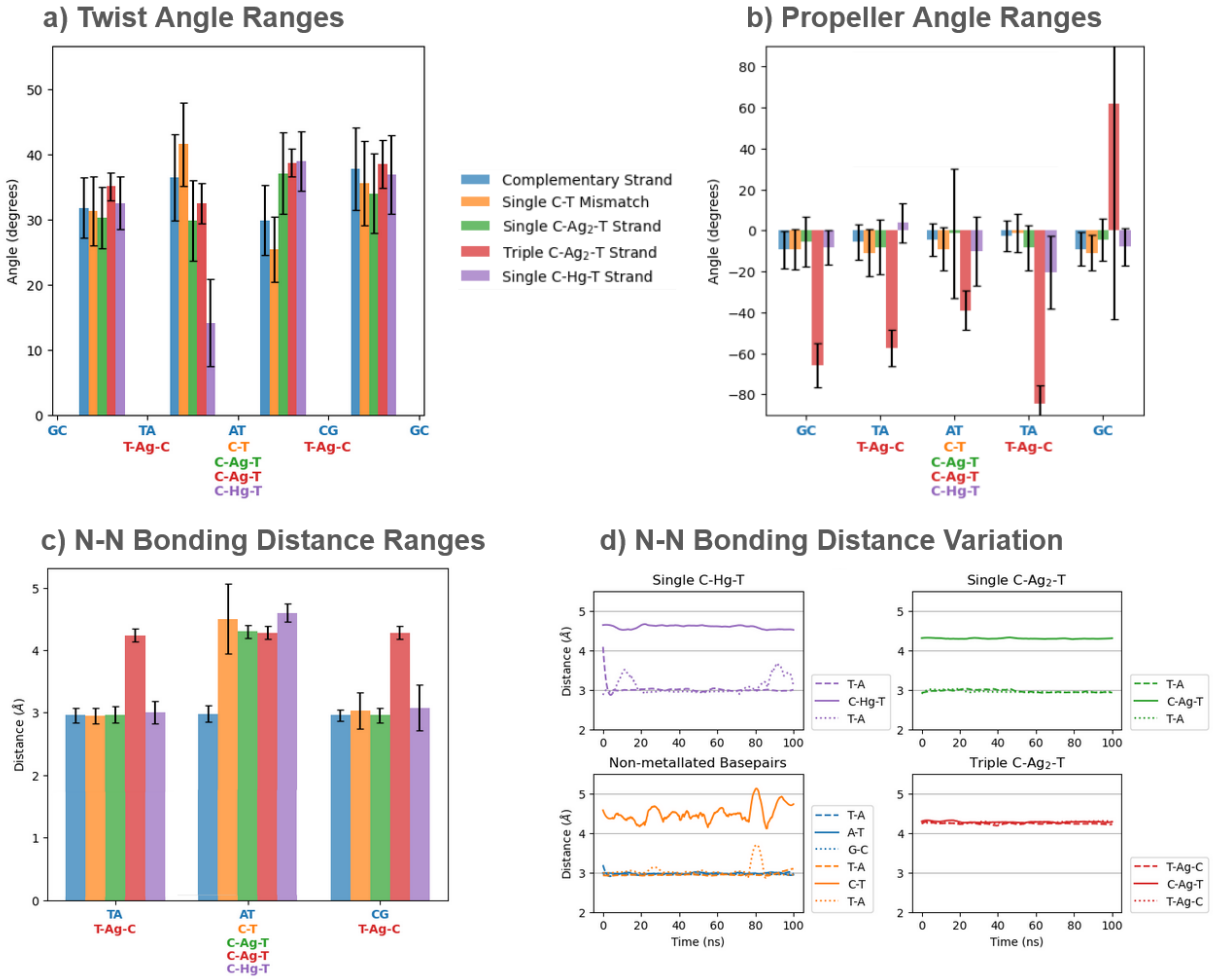


Figure 4.10 - MD statistics plotted for the seven basepair structures: (a) inter basepair twist angle, (b) basepair propeller angles, and (c) spacing between coordinating nitrogen atoms. These are shown for the complementary, single mismatch, single C-Ag_2-T, and triple C-Ag_2-T cases, with labels on X-axis color coded to match up with the five different sequences modeled. The mean is plotted, with error bars given by standard deviation across 100ns trajectory. The N-N distance variation across the MD trajectory is shown in (d), applying a Savitzky–Golay filter across a 100 ps window to smooth data, with lines color coded to match up with each case. Only data for the central five basepairs is shown.

c. Discussion

Parameterization of the C-T mismatch mmDNA structures successfully created stable force fields that can be applied to larger DNA structures. The force field parameters developed here were shown to match ab initio results with the same level of accuracy as parmBSC. This method of parameterization has been built as a patch to the existing MCPB.py tool and can be applied to any mmDNA basepair for use in the AMBER MD package, allowing one to look at stability of larger mmDNA heterostructures for nanotechnology applications.

The basepairs characterized in this work have never been parameterized for MD, but our results are comparable to similar systems. Though the force constants calculated for the C-Ag_2-T basepair did not match those previously calculated for the C-Ag-C basepair (N3-Ag bond force constant of 193 kcal mol^{-1} $Å^{-2}$), the resulting twist angle variations matches predictions for repeating C-Ag-C nanowires.[79] This was also shown experimentally in the crystal structure of a C-Ag-C nanowire (PDB ID 5IX7) which had an average twist angle of 36° across the repeating strand, even with multiple discontinuities in the DNA backbone.[64] The shift in twist angle also appeared to add instability, as shown in the RMSD plots (Figure 6). While the middle three fully-metallated basepairs showed a low RMSD between all snapshots, the cases with only a single metal basepair showed higher RMSD, likely caused by the transition from metallated to non-metallated basepairs. This is quite apparent in the full seven basepair structures, where the metallated structures have overall RMSD values higher than the mismatch structure and the complementary strand.

In our results we also observe some interesting patterns across the DNA structures. In all cases, the overall width of the basepair did not change significantly, with or without the metal. This is apparent from analysis of the distance between the C1' atoms on the sugar during the MD trajectories, which are shown in Figure 4.9. In each case, the average distance remains about 10-11 Å for all bases without significant fluctuation, only increasing to slightly above 11 Å for the C-T mismatch basepair and slightly below 10 Å for the metallated basepairs. The basepairs twist into the major groove of the DNA in the metallated cases, as shown by the 1-2 Å decrease in the C1'-C1' distance across all metallated basepairs while the base-base center of mass distance remains the same across all cases (see Figure 4.9). Additionally, our analysis of N-N distances in Figure 4.10c and Figure 4.10d reveals that metal incorporation stabilizes basepair fluctuations, where the triple metallated basepair structure shows markedly reduced RMSD compared to the single metal basepair, single mismatch, and fully complementary cases. Basepair propeller angles remain stable for the metallated basepairs, also shown in Figure 4.10b, with the triple C-Ag_2-T basepair stack significantly increasing the propeller angle between all bases compared to the single metal basepair and non-metallated cases. These observations are supported by experimentally obtained propeller angles for stacks of metallated basepairs, with angles of 34° and 20° for stacks of C-Ag-C and T-Hg-T metallated base pairs respectively compared to lower values obtained for non-metallated counterparts.[64,74] Overall, these findings establish crucial design principles for larger

heterostructures, especially when determining the flexibility and orientation of such structures in various environments.

While the stability of these mmDNA basepairs was explored thoroughly, the formation of the basepair or dissociation of the metals was not thoroughly investigated. The ab initio results suggest an asymptotic energy profile (see Figure 4.4) with an energy barrier of several eV, whereas the AMBER force field only fits to a second-order parabolic energy profile. Instead, fitting to a potential like the Morse potential may better predict these long range effects, as recently implemented for the C-Ag-C basepair.[107] For applications involving switching or modulation of mmDNA basepairs, these long range effects will be crucial for matching modeling results with experiment. Furthermore, any effects of oxidation state or spin effects were not analyzed here, as the model was applied to metals with fully closed-shell configurations and spin-orbit coupling was not modeled. These effects would likely affect the modeling results for metals with partially filled d-orbitals, where coordination may impact spin state, and chains of heavier metals, where spin-orbit coupling effects will compound. The model we have developed can serve as a starting point for such systems, with increased model complexity to be added into future work.

5. Multi-Heme Cytochrome Nanowire Modeling

While many biological processes that involve charge transport are relatively slow compared to the requirements for modern electronics, limited by ion diffusion and chemical reactions, multi-heme cytochromes such as the OmcS cytochrome found in *Geobacter sulfurreducens* and the small tetraheme cytochrome (STC) found in *Shewanella oneidensis* have been studied extensively due their high conductivity. Applications include microbial fuel cells, which take advantage of the electron transfer mechanisms to generate electricity following the seminal work from the Lovely group.[108] Within the *G. sulfurreducens* bacteria extracellular electron transport has been demonstrated to occur primarily through the OmcS protein, shown in Figure 5.1(a), participating in long-range charge transport with length scales on the order of microns.[4,109] The STC protein, depicted in Figure 5.1(b) has an intracellular role rather than acting an intercellular wire.[110] Looking at the overall electronic structure and theoretical conductivities of the two cytochromes can help explain this behavior and elucidate further properties to explore in future devices.

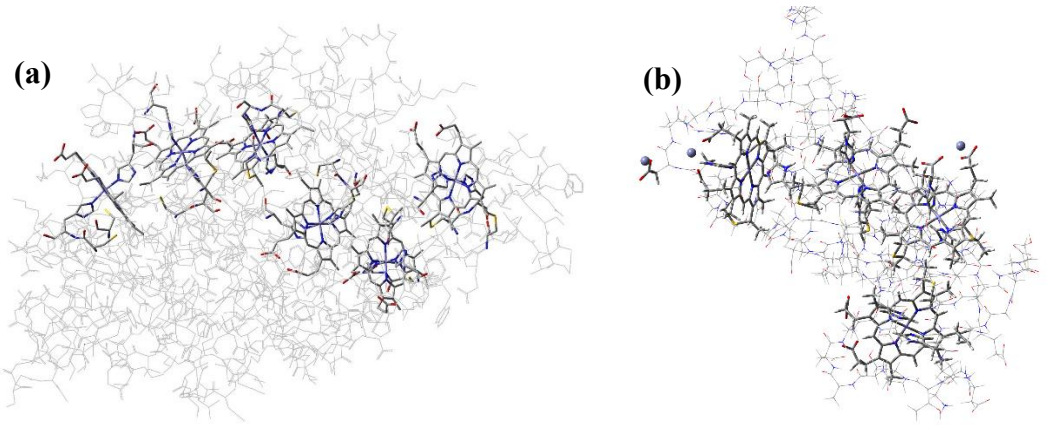


Figure 5.1 - The structures of (a) the six-heme monomer unit of the OmcS cytochrome, and (b) the STC protein with the locations of the heme porphyrin sites and bonded residues highlighted (surrounding residues in grey).

Numerous previous studies have modeled charge transport in cytochromes using Marcus theory,[5,12,111] a semi-classical electron hopping model that uses tunneling magnitude and reorganization energy to determine the kinetics of electron transfer.[13] Using this method, each heme site is modeled in its donor or acceptor state, calculating rate constants that can be solved as a system of equations to determine the current. Density functional theory (DFT) methods can be combined with other empirical methods to generate the local orbitals on each donor and acceptor site using the electronic coupling between sites to calculate a hopping rate. Other works have applied quantum transmission models to cytochrome systems, looking at ballistic (purely coherent) transport[19] and mixed coherent and incoherent (hopping) transport.[20,21] While the high conductivity of the multi-heme cytochromes may suggest a faster transport method like ballistic transport, the high fluctuations of the biomolecule seem to suggest loss of coherence over the length scales. Despite studies showing that dephasing times in cytochrome systems can approach one picosecond,[112] this corresponds to the same time scale of an electron transfer between two hemes,

and therefore loss of coherence is expected even at the heme to heme hopping length scale. While several multi-step hopping models were initially considered for this work, the energy barrier for single step electron hopping was found to be much lower than the energy barrier of intermediate reactions. In this work we explored previously studied single step incoherent hopping models, where electron transfer is limited to onsite hopping kinetics, and compared them to decoherent quantum transport models that include multiple concurrent pathways of transport, including interaction with the environment. Using open-shell DFT methods, we were able to model the system in different oxidation states, predicting overall charge transport as well as spin-dependent transport along the cytochrome.

The two cytochromes were modeled with DFT at varying length scales to generate a system Hamiltonian, which acts as a single-electron operator used to construct molecular orbitals and determine electronic band structure. Using partitioning methods to create a local orbital basis at each heme site, hopping energies were calculated for the combined system looking at the impact of oxidation state, site conformation, and spin. As a final step, a decoherence model was applied to the system and Green's functions methods were used to measure overall quantum transmission. The non-equilibrium Green's function (NEGF) method was applied to the monomer unit of the cytochrome, where interactions with the electrodes and environment were represented by self-energy functions. Using this formalism, it is possible to determine the transmission spectrum and conductance of a two terminal nanodevice constructed from the cytochrome.

a. Modeling Methods

All DFT calculations were done using Gaussian-type orbitals and the Gaussian 16 software package.[59] The B3LYP functional was used with the restricted closed-shell model used for singlet states and the restricted open-shell model used for states with net spin (non-singlet states). An exception was the free Fe ion case where HF was used. To determine the ground state spin (and subsequent spin state splitting energies), we used the unrestricted solver result as a guess for the restricted open model to increase the accuracy of the total energy with the added spin constraints.[113] The ground spin state was calculated by comparing the DFT energies of multiple spin multiplicities for the same coordinates and choosing the lowest energy case. To ensure consistency in the calculated ground spin state, DFT energies were compared within the unrestricted and restricted open models, and no discrepancies were found. The Hamiltonian was constructed based on the formula from Equation 7 using the resulting alpha and beta Fock matrices from the restricted open-shell calculation rather than the unified coupling operator matrix. A mixed orbital basis was used: the LANL2DZ basis with the effective core potential[60] for Fe atoms in the heme center and the 6-31G** Pope basis set[61] for all other atoms. To account for water solvent effect, the polarizable continuum model (PCM) was applied to account for the dielectric behavior of the aqueous environment.[62]

The free Fe ion was the only system modeled using the restricted open Hartree Fock (ROHF) method to force the full weight of the exchange term for the single atom. This was chosen due to

the lack of Coulomb attenuation included in the B3LYP functional, which leads to a significant error in the exchange term for a single atom (i.e. in the infinite limit of interatomic distance).[114] Spin-splitting energies were calculated by taking the difference in energies of each possible spin multiplicity based on 3d and 4s orbital occupation. The ferro/ferricyanide system was constructed with a starting set of coordinates that were optimized by minimizing DFT energy for both oxidation states in their ground state spin configuration.

The OmcS nanowire structure was modeled in sections ranging from one heme center to the full monomer unit using beta-carbon terminated histidine and cysteine functional groups. This method was developed in a previous work[115] based on the observation of minimal interaction with surrounding protein[111] and confirmed with DFT results showing the impact of backbone interactions on the band structure. Using the structures from this work (with initial coordinates from PDB ID 6EF8[109] and minimization of PM6 energy[44] to determine hydrogen locations), multiple oxidation states were compared, with the ground state spin multiplicity confirmed across spin multiplicities for fixed atomic coordinates. Spin alignments between heme sites were modeled by setting an initial spin constraint for each fragment and comparing the total energy of the converged SCF. For the fully reduced (Fe^{II}) state, the hemes were modeled in their deprotonated state, while all other partial and fully oxidized (Fe^{III}) states were modeled with fully protonated hemes.

Transmission calculations were conducted for smaller sections of the OmcS monomer unit to look at transport at different length scales. Changing charge and spin state at each of these length scales impacted orbital location and band energy caused by electron-electron repulsion. To model non-singlet states (with unpaired spin states), the DFT functional used a spin-unrestricted approach to generate the alpha and beta spin densities, which were constrained by the minimized spin multiplicity. The results for the minimum energy spin multiplicity were reported, and the full Hamiltonian was generated for transport calculations by separating the alpha and beta states for each atomic site. While the Green's function calculation used the combined Hamiltonian (from Equation 7) to determine the total transmission, spin-dependent transport was measured by dividing the transmission matrix into quadrants.[116] For this study, probe self-energies were applied based on previous work with other biomolecular systems ($\Gamma_{L/R} = 0.1\ eV, \Gamma_B = 0.01\ eV$).[27] Because spin-orbital coupling was not included in these calculations, the off-diagonal terms (hopping parameters between alpha spin states and beta spin states) were set to zero leading to only two reported transmission values (alpha transmission and beta transmission) without cross terms.

The model setup used for the OmcS nanowire were applied to the STC cytochrome and expanded on to look at the effects of the solvent and the backbone. In addition to the PCM treatment of the solvent used in all previous calculations, counterions were added to the STC system without solvent to simulate dry conditions. Backbone modeling was conducted used a two-layer ONIOM calculation, applying the Amber PARM96 force field and charges to the backbone

residues while still modeling the heme, Zn atoms, and coordinating residues with full ab-initio treatment.[117] Two cases were run using the two-layer setup: (i) treating all the backbone as point charges between the heme molecules and Zn atoms and (ii) including inter-heme peptide chains bonded to coordinating histidine and cysteine residues. Transmission spectra and conductances were calculated for each case to determine effects of each model variable.

b. Model Verification Results

As an initial study, the spin-separated Hamiltonian was constructed for both the free Fe ion and the smaller ferro/ferricyanide ($Fe(CN)_6$) systems. In the case of the free Fe ion, it was confirmed that high spin states were preferred (spin splitting energies ranging from 3.5-8.0eV for both oxidation states), although some of the d-orbital degeneracies were found split because of electron repulsion between occupied d-orbitals. Looking at the generated band structure, we can see that the spin-separated HOMO energies line up very closely with the experimentally determined ionization energies,[118] while the combined Fock HOMO energies do not follow this trend (See Table 2 beloq). This confirms that the spin-separated model does follow Koopman's theorem.

System	Spin Config (4s/3d)	Relative HF energy (eV)	α-HOMO (eV)	β-HOMO (eV)	Combined Fock HOMO (eV)	Reported Ionization Energy[118] (eV)

Fe Quintet	4s: ↑↓ 3d: ↑↓1111	+19.78	-6.38	**-8.20**	**-6.37**	**-7.90**
Fe(I) Sextet	4s: 1 3d: ↑↓1111	+15.13	**-15.17**	-21.79	**-7.06**	**-16.19**
Fe(II) Quintet	4s: 3d: ↑↓1111	*0.00* (reference)	-35.01	**-31.36**	**-14.75**	**-30.65**
Fe(III) Sextet	4s: 3d: 11111	-28.99	**-57.67**	-106.40	**-28.03**	**-54.80**

Table 2 - Reported spin configurations and energies from the Hartree Fock calculations for the isolated Fe atom in each oxidation state. The HOMO energies in columns 4 and 5 are calculated from the Hamiltonian in Equation 22 and the HOMO energy in column 6 is calculated from the Hamiltonian in Equation 23.

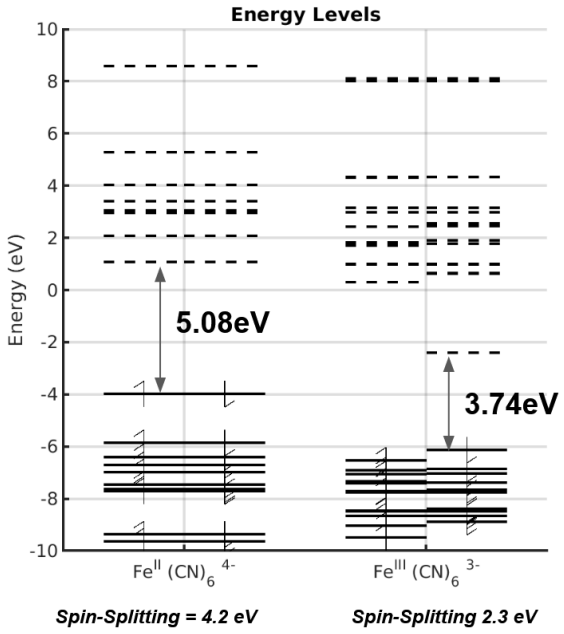


Figure 5.2 - Energy level diagrams for ferrocyanide (left) and ferricyanide (right) showing the calculated band gaps and high-spin/low-spin splitting energies. The bandgaps of 5.08 and 3.74 eV correspond to photon wavelengths of 243 and 330 nm respectively.

As a second test, the energy levels of the coordinated ferricyanide (oxidized state) and ferrocyanide (reduced state) ion were modeled using the spin-separated Hamiltonian. The ferro/ferricyanide system was chosen for comparison due to similarity with the heme system, which also has octahedral coordination with similarly spaced organic ligands (2-2.3 Angstroms in the heme case vs 1.96 Angstroms in the ferro/ferricyanide case). Upon coordination in the ferro/ferricyanide case, any d-orbital degeneracy was broken, and the low-spin state was energetically preferred, as predicted by ligand field theory and confirmed by DFT calculations. The resulting band diagrams (and spin-splitting energies) are shown in Figure 5.2 and the band gaps can be directly compared to absorption peaks that have been identified for each oxidation

state, ferrocyanide vs ferricyanide, which are found at wavelengths ranges 205-235nm and 310-340nm respectively.[119] This result serves as experimental validation of the band gap calculated from the constrained unrestricted approach and as further confirmation of spin separate treatment for transport modeling.

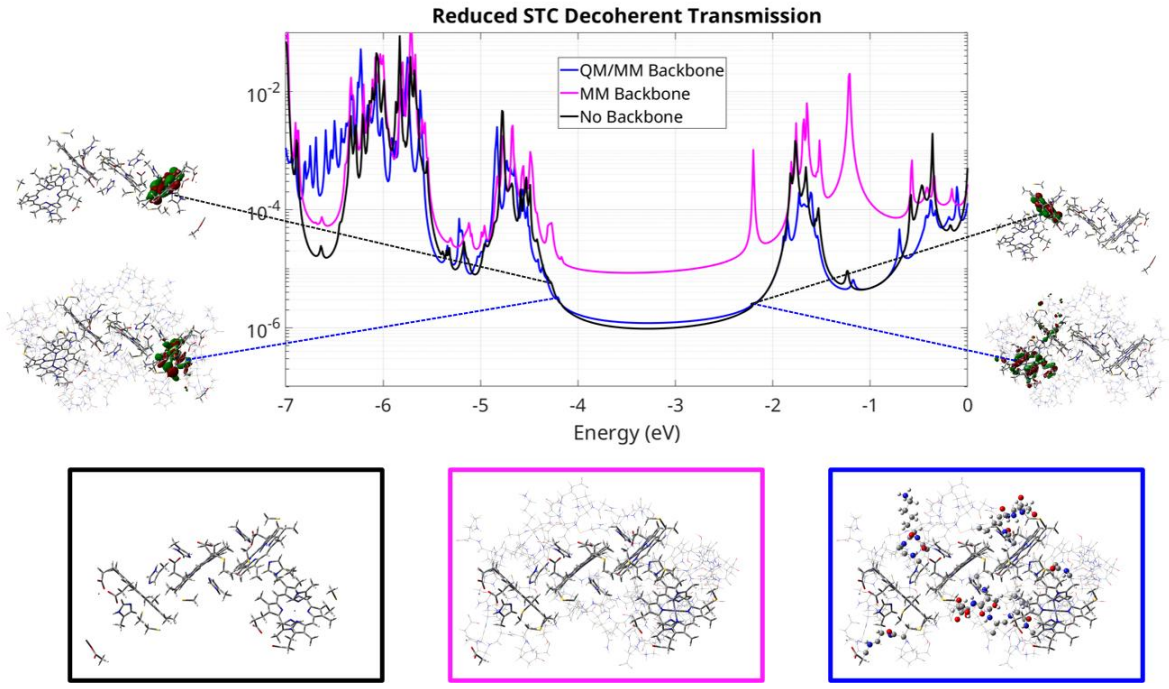


Figure 5.3 – Decoherent transmission comparison for various levels of modeling applied to the small tetraheme cytochrome, with HOMO/LUMO wavefunction orbital plotted for the no backbone and mixed QM/MM backbone cases (with cutoff iso values $|\psi| < 0.02$)

To look at the effects of the backbone, the decoherent transmission was compared for several modeled cases using the STC protein. The fully ab-initio model was compared to two QM/MM models modeling the backbone as point charges and then applying QM to connecting backbone residues. The resulting transmission curves are overlayed in Figure 5.3, along with HOMO/LUMO

orbitals plotted for selected cases. As can be seen from the results, using only point charges for the backbone significantly increased the decoherent transmission in the band gap, whereas the mixed backbone almost completely matched the coordinated heme result. For this reason, in all further studies the backbone was not modeled for either cytochrome, only the hemes with bonded residues terminated at beta carbons.

c. Cytochrome Modeling Results

Ligand field theory has shown that while the uncoordinated heme prefers a high spin state, axial histidine coordination on both sides of the heme shifts the point group from D4h to Oh and the ground state prefers the lower spin state for both reduced and oxidized hemes. This was confirmed for the single hemes coordinated with histidines on both sides -- the reduced Fe (II) system shows energetic preference to the singlet state and oxidized Fe (III) system shows energetic preference to the doublet state. High spin ground states only occur with uncoordinated or single-sided coordinated hemes, as shown previously[120] and confirmed with our model. Orbitals were found to be more delocalized in the charged Fe (III) oxidation state, with the band gap lowering from 2.9eV to 2.0eV. A reorganization energy of 0.11 eV was calculated between the two charge states, with a slight band structure shift occurring in the reduced Fe (II) charge state, as shown in Figure 5.4. To check the alignment of spin states, two-heme pairs were modeled in parallel and perpendicular configurations, and spin splitting energies were calculated. In the two-heme case the low spin ground state was confirmed for each individual heme and spin alignment was

energetically favorable between heme sites both in the parallel and perpendicular conformations. This spin configuration was applied to all larger fragments as the ground state electronic structure.

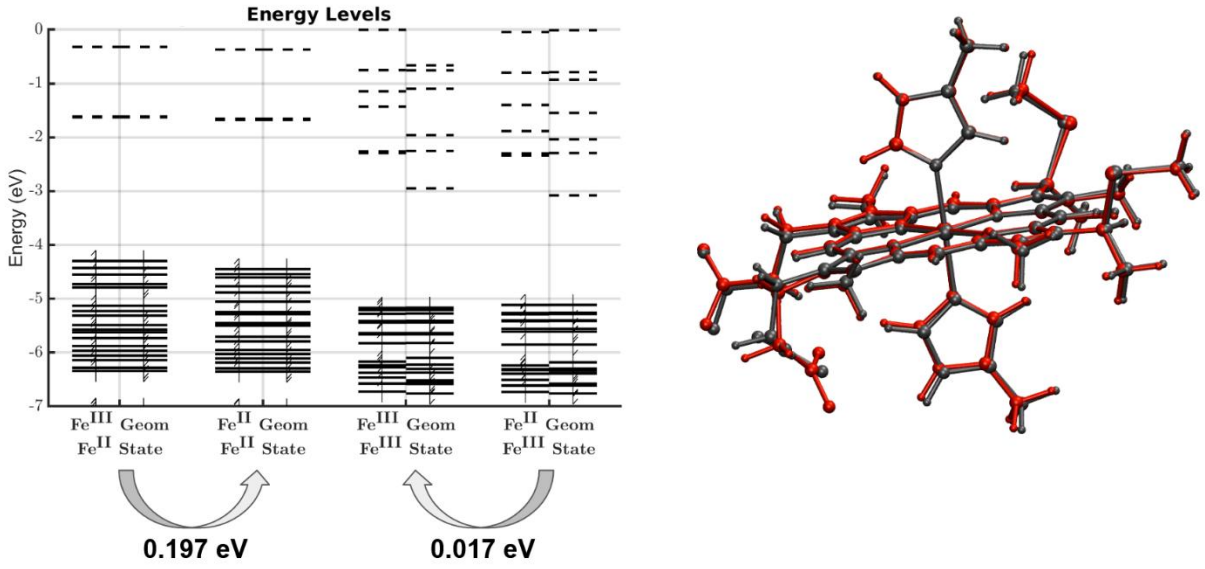


Figure 5.4 - Reorganization impact on the deprotonated heme structure, showing the shift in band structure during reorganization on the left and the physical coordinate changes in the right, where the red structure is in the oxidized state and the grey structure is in the reduced state. DFT energy differences are calculated between geometries at the same oxidation state.

Hopping energies were calculated for all four charge states of the three-heme subunit to compare all pathways of charge transfer between heme sites within the OmcS cytochrome. The three-heme sub-unit was chosen as a model system because of the presence of both parallel and perpendicular conformations, allowing calculation of nearest neighbor hopping energies for all configurations. In addition, the four oxidation states represent four charge transfer methods: (a) fully reduced, HOMO-HOMO transitions, Q=-6; (b) hole transport, HOMO-LUMO transitions with oxidized central heme, Q=+1; (c) electron transport, HOMO-LUMO transitions with reduced

central heme, Q=+2; and (d) fully oxidized, HOMO-HOMO transitions, Q=+3. This means that case (b) was calculated as electron transfer from the outer hemes to the central heme and case (c) was calculated as electron transfer from the central heme to the outer hemes. These cases (a-d) are depicted visually on the three hemes in Figure 5.5 below.

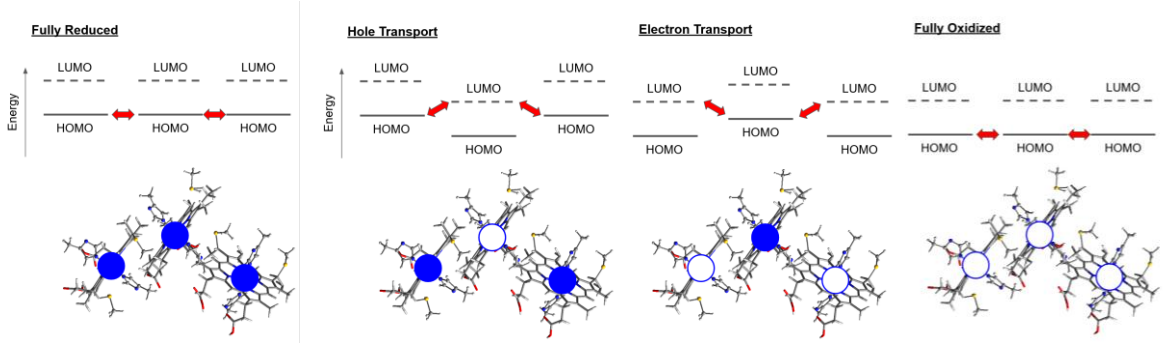


Figure 5.5 - A visualization of the four scenarios used for hopping energy calculations applied to the three heme sub-unit.

While we do not fully include effects of orbital reorganization between hemes during electron transfer, this model includes the hopping energy shifts during the transition state, where an electron would likely travel from the HOMO state of a reduced heme site to the LUMO state of the unoccupied heme site. Furthermore, this includes the charging energy that causes the local orbitals to shift when an electron is added or removed, depending on the direction of electron transport. For this reason, we label "hole transport" as the transition of the +1 (oxidized) charge state to the neutral (reduced) neighboring hemes, and this is measured by looking at the HOMO to LUMO couplings of the neighboring hemes to the central heme respectively. In the case of "electron transport" we look at the transition of the neutral (reduced) charge state to the charged (oxidized) neighboring hemes, specifically looking at the HOMO to LUMO couplings of the central heme to

the neighboring hemes respectively. These electron trajectories are shown in Figure 5.5, with spin polarization separating this calculation into two parts, one set of coupling values for alpha spin and one for beta spin.

The charge and spin populations were confirmed for each of these four states by using the full natural bond orbital (NBO) method for analysis.[121] We have reported the NBO spin and NBO charge of the Fe ions in each of the three hemes (labeled 1-3 matching up with the hemes left to right shown in each case in Figure 5.5) as wells as the sum of the NBO charge on the carboxylate groups in Table 3 below. Note that NBO population data was collected for each system modeled to confirm the locations of spin and charge across the system.

		Heme 1			**Heme 2**			**Heme 3**		
Total Charge	COO^- /COOH	Fe Charge	Fe Spin	2xCOO(H) Charge	Fe Charge	Fe Spin	2xCOO(H) Charge	Fe Charge	Fe Spin	2xCOO(H) Charge
Q=-6	COO^-	0.241	0.000	-1.719	0.27	0.000	-1.732	0.243	0.000	-1.733
Q=+1	COOH	0.509	0.000	-0.052	0.822	**0.889**	-0.004	0.531	0.000	-0.057
Q=+2	COOH	0.520	**0.875**	0.052	0.274	0.000	0.060	0.533	**0.891**	0.050
Q=+3	COOH	0.520	**0.875**	0.052	0.558	**0.889**	0.075	0.533	**0.891**	0.051

Table 3 - NBO population data reported for the Fe and carboxylate groups on each heme in the three-heme subunit for each of the oxidation states modeled.

We determined the hopping energies between perpendicular hemes were an order of magnitude smaller than parallel hemes for OmcS in all cases, matching up with previous calculations[115] as depicted in Figure 5.6 and listed in Table 4. The hopping energies for the oxidized case have a

significant spin dependency, resulting from the spin alignment between the hemes discussed earlier. The ratios of the alpha and beta charge transfer energies were plotted in the right plot of Figure 5.6, showing the spin dependency especially in the Q=+1 hole-transfer case favoring the beta (valence) spin state.

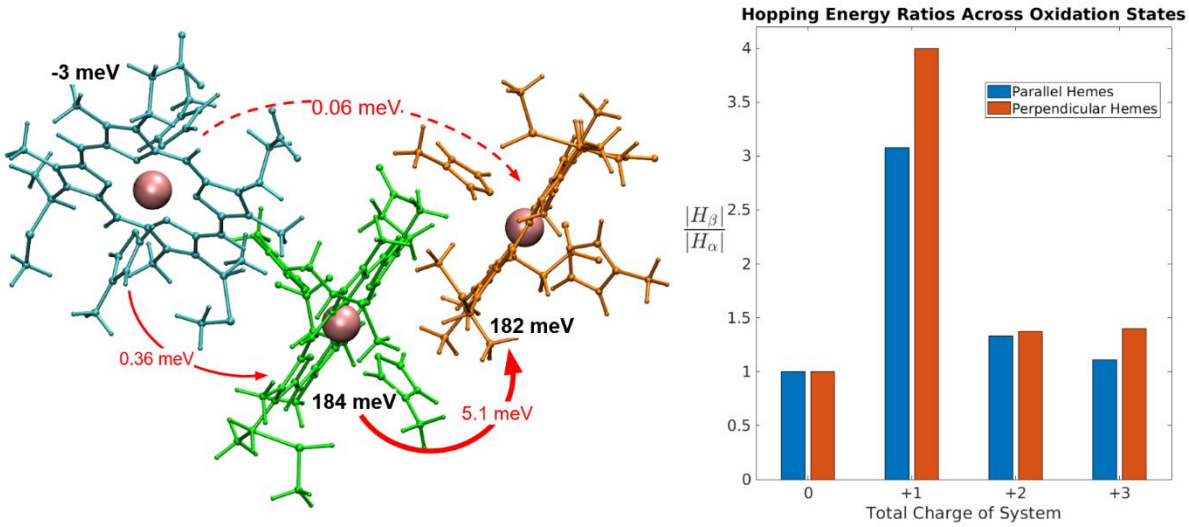


Figure 5.6 - Modeling results from the three-heme subunit showing the relative onsite and hopping energies for the reduced case on the left and the spin dependence of hopping energies on the right.

Hopping Energy Calculation	Q =-6 (HOMO-HOMO)	Q=+1 (HOMO-LUMO)	Q=+2 (LUMO-HOMO)	Q=+3 (HOMO-HOMO)
α-Parallel	5.1 meV	2.6 meV	4.2 meV	10.7 meV
β-Parallel	5.1 meV	8.0 meV	5.6 meV	11.9 meV
α-Perpendicular	0.4 meV	0.1 meV	0.8 meV	0.5 meV
β-Perpendicular	0.4 meV	0.4 meV	1.1 meV	0.7 meV

Table 4 - Calculated hopping energies for each of the oxidation states and both conformations in the three-heme subunit

Hopping and decoherent were applied to the full OmcS monomer unit to generate a system Hamiltonian. For this six-heme system, four cases were modeled: reduced Fe^{II} vs oxidized Fe^{III} and deprotonated COO^- vs protonated COOH on the propionate groups. An analysis of incoherent hoping was conducted on the fully reduced (Fe^{II} and COO^-) and the fully oxidized (Fe^{III} and COOH) states, with HOMO-HOMO hopping energies plotted in Figure 5.7a. These results show a drastic increase in hopping energy for the oxidized case as well as some spin dependency. This is especially notable for the hopping energy between hemes 3 and 4, which have a perpendicular conformation, showing the importance of modeling electron delocalization across hemes in the nanowire. A fuller picture of electron transport was constructed by using Green's functions to model all orbital interactions. This was used to calculate local density of states for the fully oxidized case as a function of energy for each heme, looking at both the total electron density and spin density ratio as shown in Figure 5.7c. These plots show current pathways at Fermi energies near the HOMO and LUMO bands with high spin polarization in the LUMO band. This was confirmed by plotting the spin-dependent transmission in Figure 5.7d, which shows up to 100x difference in decoherent transmission between alpha and beta spins at higher Fermi energies. The decoherent transmission was compared for the deprotonated COO^- cases as a function of Fermi Energy and plotted in Figure 5.7b to compare with experimental results at neutral pH. The calculated transmission at energies between -5 and -4.5eV are significantly higher for the reduced case despite low HOMO-HOMO hopping energies, most likely related to the shift in HOMO energy between the oxidation states. In addition, the transmission at the HOMO energy for the

reduced case is about 4x higher than for the oxidized case. Both observations from the decoherent model qualitatively match up with recent experimental results.[5] A separate calculation confirms that the shift in transmission curve is caused primarily by the change in oxidation state of the Fe center, with only a small contribution from reduction of the carboxylate groups on the heme, see Figure 5.8 - A comparison of decoherent transmission of the monomer unit of the OmcS cytochrome looking at four Figure 5.8 for a direct comparison of the results for these four cases.

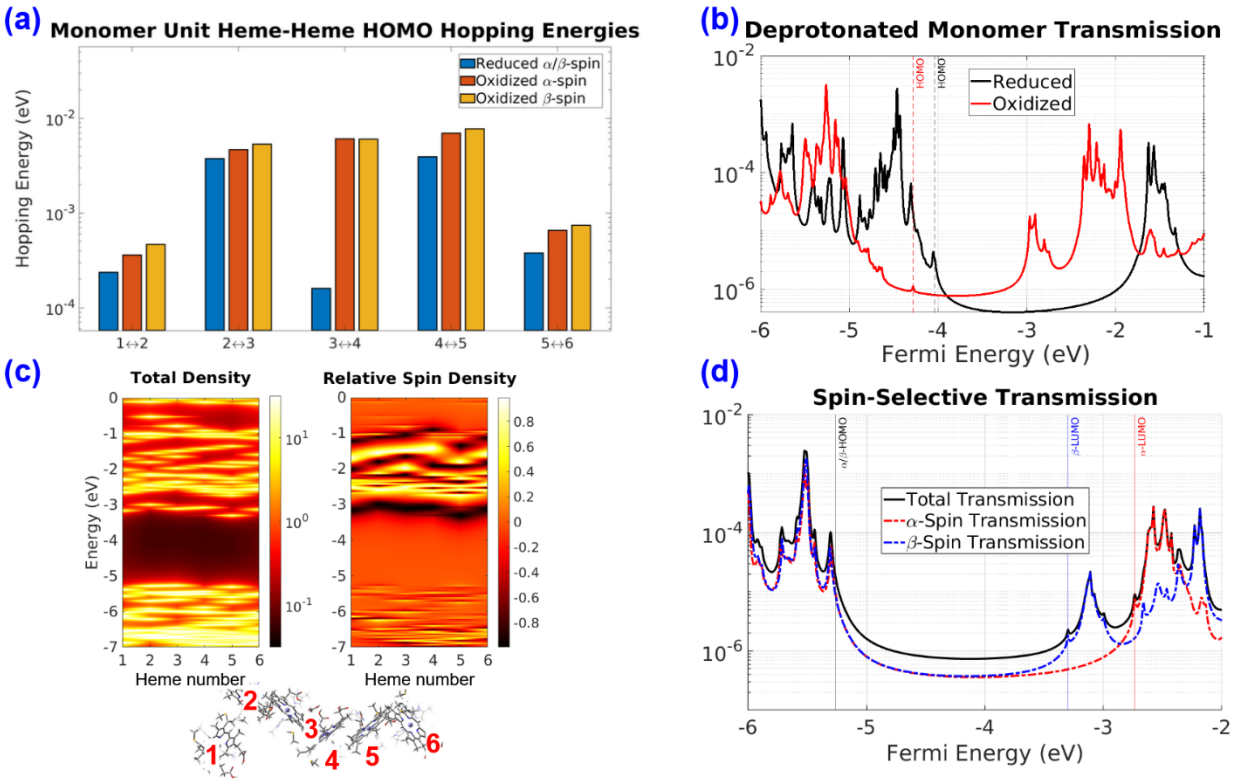


Figure 5.7 – Electron transport across OmcS monomer with (a) comparison of HOMO-HOMO hopping energies for the fully oxidized and fully reduced cases, (b) comparison of quantum transmission of the nanowire in the deprotonated (COO^-) oxidized and reduced cases, (c) the total density of states across the fully oxidized monomer,

segmented by hemes labeled in the diagram, as well as the relative spin density, given by $(\rho_\alpha - \rho_\beta)/\rho_{tot}$*, and (d) the spin-dependency of transmission for the fully oxidized monomer.*

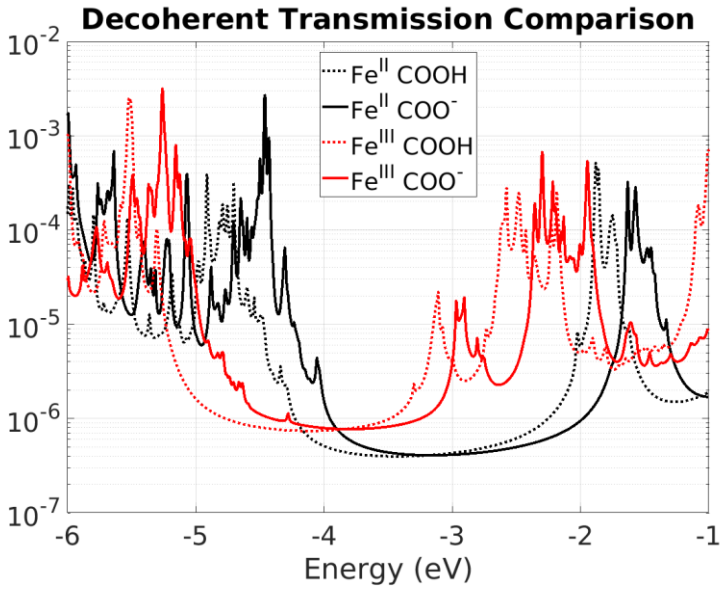


Figure 5.8 - A comparison of decoherent transmission of the monomer unit of the OmcS cytochrome looking at four different oxidation states.

All oxidation and spin states modeled for each section of the OmcS cytochrome modeled are listed in the table below, with total DFT ground state energies reported for each oxidation state:

System/ Oxidation State	Total System Charge	Restricted Closed Shell		Restricted Open Shell		Unrestricted Open Shell		Total Ground State Energy (Hartree)
		Spin Mult	Energy (eV)	Spin Mult	Energy (eV)	Spin Mult	Energy (eV)	
Single heme Fe^{II} (COO-)	-2	**1**	**0.00**	**1**	**0.00**	**1**	**0.00**	-2588.96713945
				3	+1.43eV	3	+1.15	
				5	+3.00eV	5	+1.09	
Single heme Fe^{III} (COOH)	1	--	--	**2**	**0.00**	2	-0.04	-2589.77368923
				4	+5.63	4	+1.81	
				6	+3.89	6	+3.81	
Two Hemes Parallel (Fe^{II}, Fe^{II}) (COOH)	0	**1**	**0.00**	**1**	**0.00**	**1**	**0.00**	-5183.05252681
						3	+0.87	
						5	+28.16	

Two Hemes Parallel (Fe^{III}, Fe^{III}) (COOH)	+2	1	+1.79	1	+1.79	1	+1.79	-5183.11853602
				3	**0.00**	3	-0.10	
						5	+26.28	
Two Hemes Perpendicular (Fe^{II}, Fe^{II}) (COOH)	0	**1**	**0.00**	**1**	**0.00**	**1**	**0.00**	-5182.00246566
						3	+2.44	
						5	+30.73	
Two Hemes Perpendicular (Fe^{III}, Fe^{III}) (COOH)	+2	1	+2.00	1	+2.00	1	+2.00	-5182.06869594
				3	**0.00**	3	-0.20	
						5	+23.34	
Three Hemes (Fe^{II} X 3) (COO-)	-6	**1**	**0.00**	**--**	**--**	**--**	**--**	-7771.14137518
Three Hemes (Fe^{II}, Fe^{III}, Fe^{II}) (COOH)	+1	**--**	**--**	**2**	**0.00**	2	-0.05	-7773.87884690
Three Hemes (Fe^{III}, Fe^{II}, Fe^{III}) (COOH)	+2	**--**	**--**	1	+1.78	1	-0.1	-7773.74846307
				3	**0.00**	3	-0.09	
Three Hemes (Fe^{III} x 3) (COOH)	+3	**--**	**--**	**4**	**0.00**	4	-0.14	-7773.60702336
Six hemes (Fe^{II} X 6) (COO-)	-12	**1**	**0.00**	**--**	**--**	**--**	**--**	-15546.4030759
Six hemes (Fe^{III} X 6) (COOH)	+6	1	+5.68	1	+5.68	1	+5.68	-15551.3538296
				7	**0.00**	7	-0.23	
						13	+0.84	
Six hemes (Fe^{III} X 6) (COO-)	-6	**--**	**--**	7	**0.00**	7	-0.28	**-15545.6319136**

Table S1 –

Applying the same model to the STC cytochrome, we find the same spin dependence result for the oxidized case. Using equation 33 we can use the transmission to calculate zero bias conductance at room temperature, which allows use to compare the results to experimentally measured conductance values[122] as shown in Figure 5.9 – Figure 5.9A. For contact Fermi energies in the HOMO and LUMO bands, the conductance measurements match up with experimental values in both the oxidized and reduced case. The effect of solvation on the orbitals on the reduced case is shown in in Figure 5.9B, with counterions added for the vacuum case to represent charge neutrality outside of a solvent and to ensure convergence of the DFT calculations. The LUMO orbitals from the counter ions exist right above the HOMO level of the structure, and significantly increase the decoherent transport as shown in Figure 5.9C, though these localized orbitals don't hybridize to create a single current pathway and the transmission remains relatively flat across the HOMO LUMO gap.

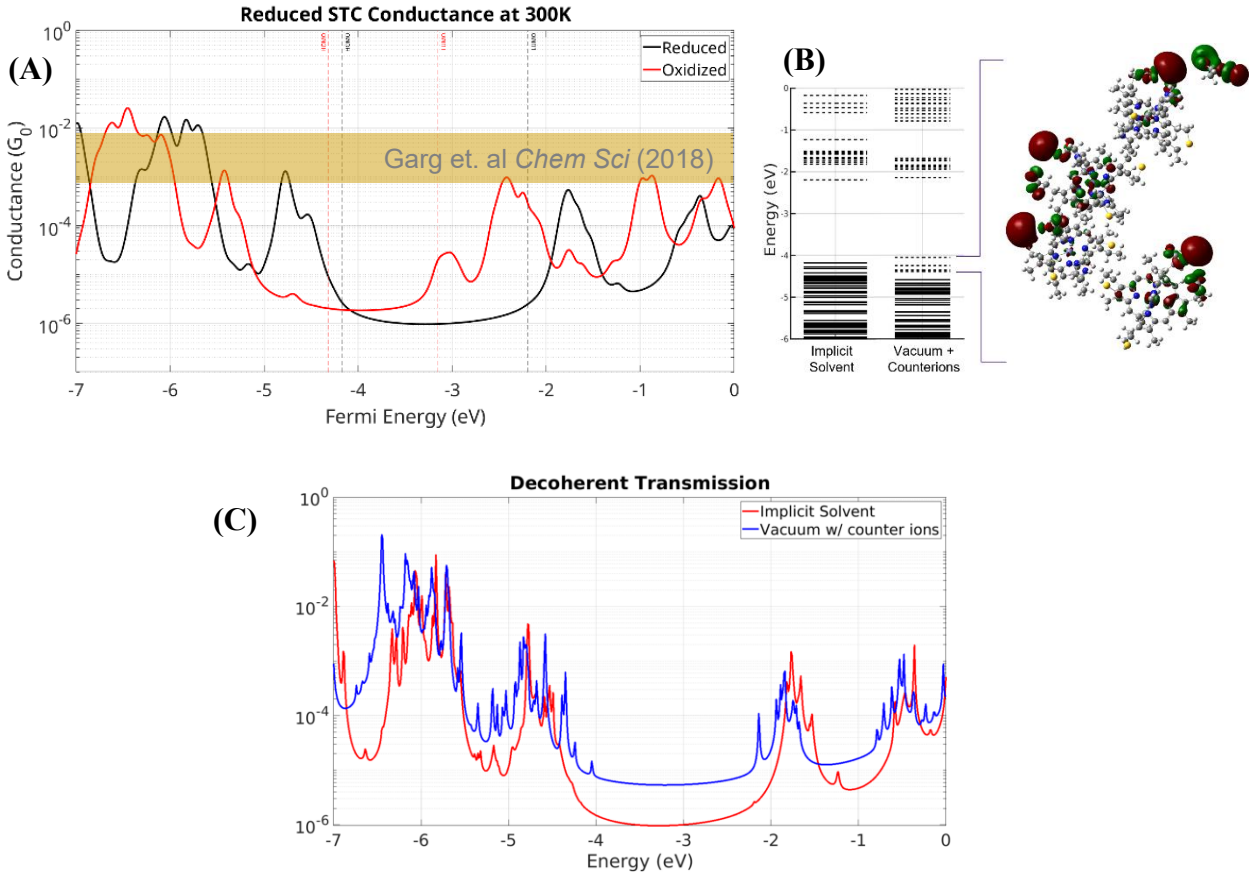


Figure 5.9 – Decoherent transport results for the STC cytochrome: (A) the difference in zero-bias conductance for the reduced and oxidized case as a function of Fermi energy, (B) the effect of solvation on orbital energies in the structure, with isosurface plots shown for each of the counterion LUMO orbitals ($|\psi|$=0.02), and (C) a comparison of decoherent transmission for the solvent and vacuum cases as a function of electron energy.

d. Discussion

To understand electron transport in multi-heme cytochromes, ab-initio methods to generate molecular orbitals of the full multi-heme structure. This allows us to consider all mechanisms of charge transfer across the protein, rather than applying constraints to specific donor and acceptor sites. Hopping energy calculations did not include any kind of nuclear coordinate shift, although

reorganization was found to not significantly affect the orbital energies for the single heme case. Although the hopping energy between heme sites in the cytochrome is smaller than those of other biological systems, like DNA, we have shown that orbital delocalization plays a key role in transport, with band formation in the oxidized state creating a significant increase in tunneling rates near the HOMO energy. The open-shell methods used to generate a spin-separated Hamiltonian can be applied to any system to develop an electron transport model, especially when considering interactions between many orbitals. While unrestricted models are computationally efficient and tend to be more accurate for determining ground state energies due to fewer constraints on the system, the resulting molecular orbitals can break spin symmetry. We propose using the restricted open model to generate the Hamiltonian and include multiple concurrent pathways across multiple oxidation states as well as spin-dependent electron transport. We also found that the implementation of the restricted open-shell solver greatly reduced spin contamination for all cases and shifted the size and magnitude of the transmission peaks. Green's function methods allow us to include all interactions of orbitals for modeling electron transport and consider a mix of different charge states. We were able to use the NEGF formulation to model transport across different oxidation states and confirm the experimental results, which show that the oxidation of the OmcS reduces conductivity. While hopping energies between neighboring sites were calculated to be higher in the fully oxidized case, the overall transmission is higher for the reduced case at lower Fermi energies, which can be attributed to the overall shift in orbital energies.

Our decoherent transport model does not explicitly include re-organization or any nuclear coordinate shift, an assumption supported by the single heme reorganization calculations. Instead, energy fluctuations caused by interactions with the environment were accounted for by the decoherence probes applied using a model that has been developed in previous works.[27,28] Even to include non-adiabatic effects like reorganization energy there are several intermediate steps that must be accounted for besides the total shift in energy of the system. Marcus theory does not consider the full pathway of charge transfer between the donor and acceptor, which includes a shift in energy between donor and acceptor orbitals as well as charging effects that occur during the transition. Looking at the case of the single coordinated heme site there is a local charging and discharging energy from the polarization of orbitals during charge transfer, which can be calculated by looking at the shift in energies of the LUMO orbital as it becomes occupied.[123] We have determined these energies to be significant (0.6-0.7eV for each heme site) when compared to the total free energy of oxidation (3.7-4.2 eV for each heme site, see Table 5 below). To include such effects in the model, we propose using a self-consistent NEGF formulation and looking at wavefunction dynamics during electron transfer. This would need to be accomplished on a smaller scale model, as the computation costs of such a calculation require numerous DFT calculations to converge the electron density at each time step.

Table 5 - DFT energies for each oxidation state and minimized geometry as well as the calculated reorganization energies.

	DFT Energy (Hartree)					
State	Fe^{II}	Fe^{III}			Averaged Reorg	ΔG_{ox} (eV)

Geom	Fe^{III}	Fe^{II}	Fe^{III}	Fe^{II}	Fe(II) Reorg (eV)	Fe(III) Reorg (eV)	Energy (eV)	
(a) Constrained, COOH	-2589.8928	-2589.8948	-2589.7385	-2589.7382	0.055	0.008	0.032	4.253
(b) Constrained, COO-	-2588.9436	-2588.9464	-2588.7975	-2588.7965	0.075	0.026	0.051	4.052
(c) Unconstrained, COOH	-2589.9101	-2589.9126	-2589.7737	-2589.7705	0.068	0.086	0.077	3.781
(d) Unconstrained, COO-	-2588.9599	-2588.9671	-2588.8317	-2588.8311	0.197	0.017	0.107	3.685

The consideration of spin was crucial for modeling oxidation states because removal of an electron from a pair inherently generates spin. This is especially true for an organometallic system like heme where the Fe center can exist in several stable oxidation states. By applying a collinear spin model, we have shown the impact spin can have on electron transport in the system even with higher order effects ignored. While these effects would not significantly contribute to the Hamiltonian due to the low electron transmission rates and low atomic weight of all atoms, this model does not account for effects like spin-orbit coupling as well as other spin-dependent effects like chiral-induced spin selectivity, partial orbital occupation from self-consistent calculations, and spin transitions caused by decoherence. Our results demonstrate significant spin-dependent transport, especially in the valence band, and we suspect that considering these higher-order effects will accentuate these spin-dependent effects.

We suggest further experiments to measure the spin-dependence of electron transport, especially to develop spintronics based on these protein nanowires. Our models consider ambient conditions, both with the choice to model with an implicit solvent model as well as the temperature of 300K used to determine the orbital occupation (any degeneracies caused by thermal energy). Therefore, we would recommend an experimental setup in these conditions using ferromagnetic contacts that would allow one to polarize the electron spin entering the cytochrome to measure magnetoresistance. Such a setup would be similar to that of Xiong et al that was used to look at spin filtering in organic materials.[124] Measuring the magnetoresistance and hysteresis effects in this system is paramount to incorporation in future nanodevices.

6. Geometric and Non-Collinear Spin Effects in Cytochrome Electron Transport

Electron transport in multi-heme cytochromes has been modeled extensively using or combining quantum-based transport models and semi-classical electron transfer models. While some studies have shown that pi-pi stacking in the protein backbone can contribute to electron transport,[20] there is growing consensus that transport occurs primarily between heme sites,[109,115,122] though the predicted mode of transport varies by system and level of theory.[5,21,125,126] Recent experimental results indicate that spin transport can be measured in cytochrome systems,[127,128] though supporting theory and models describing spin transport remain limited.[129,130] In Chapter 5, a spin dependent model for transport was introduced, governed by the central Fe atom in each heme porphyrin, shown in Figure 6.1, which can exist in multiple oxidation states.[129] Building upon this model, this work will look at other sources of spin-dependence, both within the heme cofactor, looking at higher spin states, as well as intrinsic to the multi-heme system, using non-collinear modeling to look at spin mixing across the OmcS cytochrome.

While consideration of spin in cytochrome electron transport models has been limited, the coordinated Fe in each heme creates inherent spin dependence. This is because the oxidized Fe^{III} state must contain an unpaired electron, regardless of coordination, which creates spin dependence in the local electron density. This spin dependence can impact hopping, changing the energy barrier

of electron transfer between a donor heme site, in the Fe^{II} state, and an acceptor heme, in the Fe^{III} state. For quantum-based models, the unpaired spin shifts all corresponding occupied and valence orbitals, which can generate spin dependence in the quantum transmission across certain energy ranges. Current spin models applied to cytochrome systems are limited to unperturbed electronic states and collinear models, which fall short of describing the most recent experimental results[128] and limit potential applications of the cytochrome system. Expanding these models to include transitions between multiple electronic states and non-collinear spin effects can elucidate new connections between spin state and electron transport.

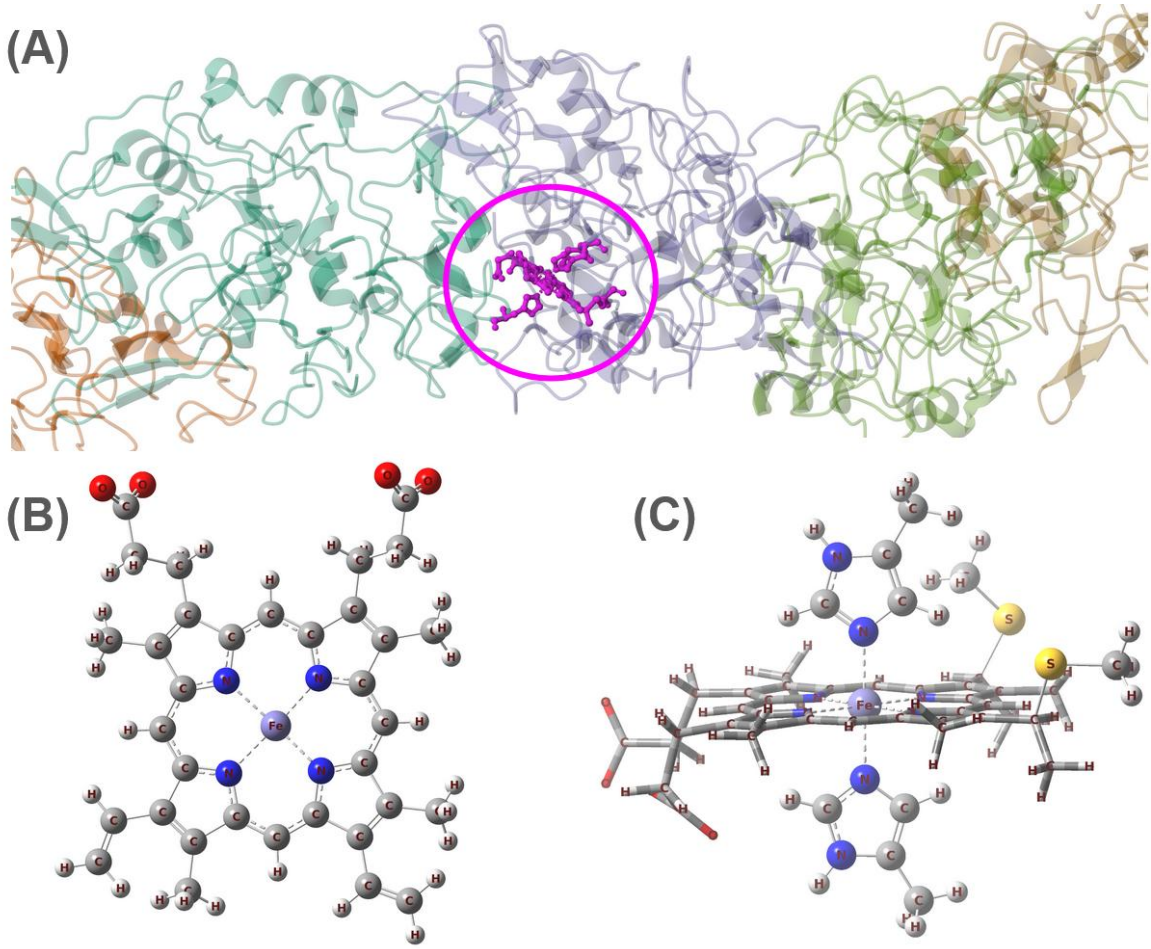


Figure 6.1 - The structure of the coordinated heme showing (A) where the heme residue is located in the OmcS cytochrome (PDB ID 6EF8), (B) the labeled structure of the isolated heme-B molecule and (C) the structure of the coordinated heme-B molecule with beta carbon terminated histidine and cysteine residues

a. Spin dependent Hopping

Electron transfer in cytochromes has been modeled for decades using Marcus Theory to describe kinetic effects, ignoring spin effects in most cases.[12,111,14] This model uses an electron hopping rate formula given by

$$k_{ET} = \frac{2\pi}{\hbar}|H_{AB}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left[-\frac{(\lambda + \Delta G^\circ)^2}{4\,\lambda k_B T}\right] \quad (39)$$

where H_{AB} is the hopping energy between the donor and acceptor sites, λ is the reorganization energy, and ΔG° is the free energy difference between sites.[13] Spin dependence can be added directly by considering the hopping energy for the electron transitions modeled in Figure 6.2, as implemented in Chapter 5.[129]

In Chapter 5, we applied Marcus theory to the cytochrome system, two limiting cases emerge in the fully reduced and fully oxidized cases. In the fully reduced case, charge transfer is governed by hole transfer (that is, the acceptor Fe^{III} site is the minority charge carrier) and the oxidized case is governed by electron transfer (the donor Fe^{II} site is the minority charge carrier). Spin dependence emerges for these two cases as shown in Figure 6.2, with the alpha spin (majority spin carrier) depicted in red and the beta spin (minority spin carrier) shown in blue. In this mechanism, transfer of the beta spin (shown in blue in the figure) can occur between low spin states with zero net spin difference (doublet to doublet transition), whereas charge transfer of the alpha spin increases the total spin (singlet to triplet transition). This suggests that alpha spin hopping in this model is not possible without a spin-forbidden transition. While this can be conclusive by itself, suggesting that beta spin electron hopping is favorable, the specific mechanism of alpha spin transport requires investigation of these higher spin states. Relaxing atomic constraints we can determine any spin transition governed by atom reorganization, which can indicate the potential for spin transport across these higher spin states.

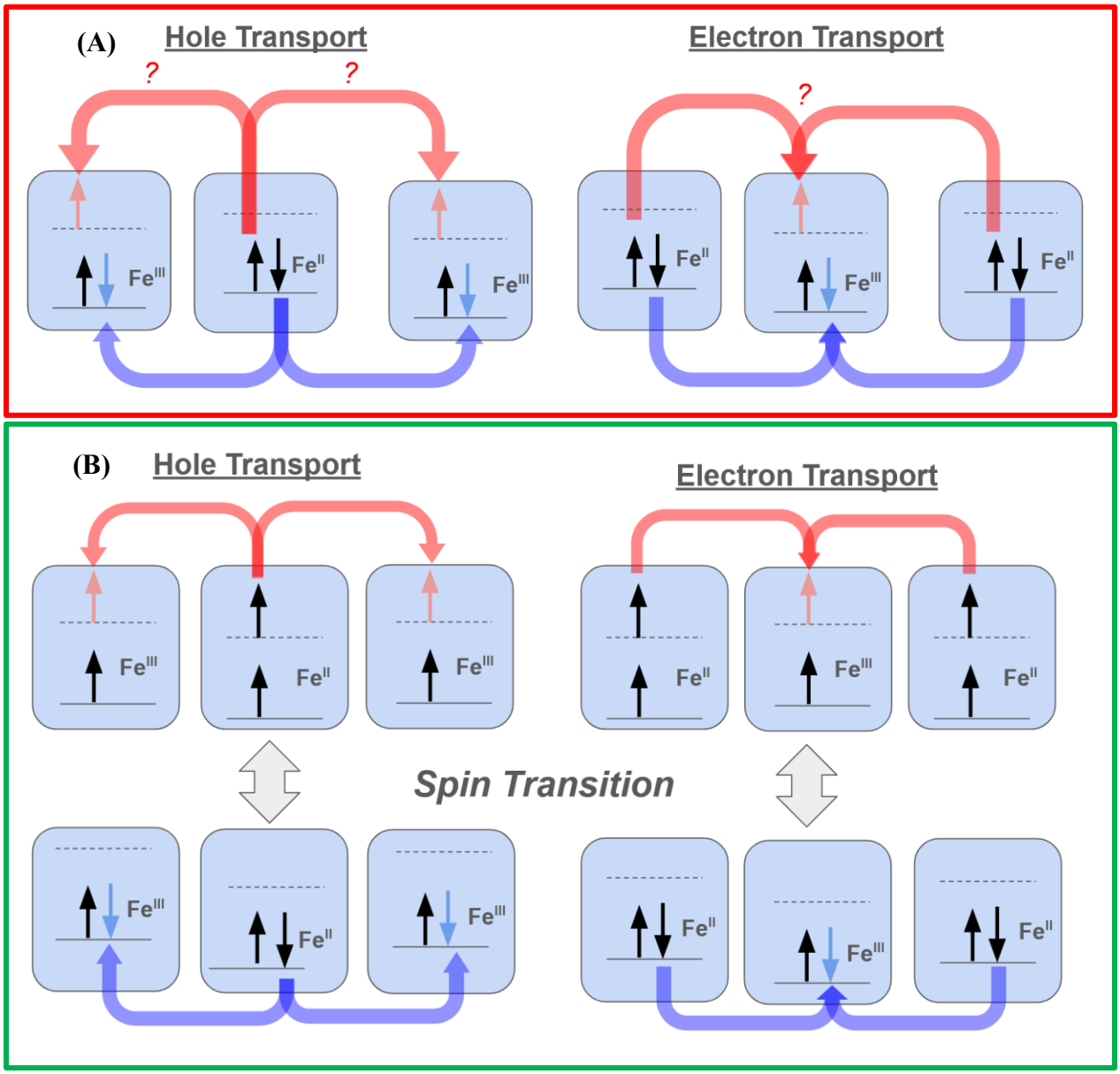


Figure 6.2 - The models used to simulate hole transport and electron transport for the three heme structures. The HOMO-LUMO transition is spin dependent: red represents alpha spin and blue represents beta spin, with the light-color electrons showing the final location after charge transfer. Previous work has relied on the hopping model shown in (A), where a spin-forbidden transition is required for electron transfer. In this work we explore model (B), where alpha spin hopping occurs only after a spin transition.

Applying energy minimization to these states, the atomic coordinates shift significantly from the original coordinated geometry both for the oxidized and reduced heme at higher spin states. For both cases, the low spin state was the closest to the cryoEM structure[109] with the histidines and heme molecule forming a highly symmetric octahedral coordinated Fe center. This matches the predictions of crystal field theory, where a low spin state indicates a strong ligand interaction in an octahedral field. Imposing a higher spin boundary condition forced the symmetry to be broken by the coordinating histidine ligands. In the reduced case, one coordinating histidine leaves the coordinating position, causing a shift to a local C4v symmetry that favors the quintet (high spin) state. By contrast, in the oxidized case, both histidines move symmetrically away from the Fe, creating a coordinating field with D4h symmetry. These results match up with predictions from crystal field theory, with the d-orbitals splitting into separate groups as depicted in Figure 6.3 along with the optimized geometries from the DFT energy minimizations.

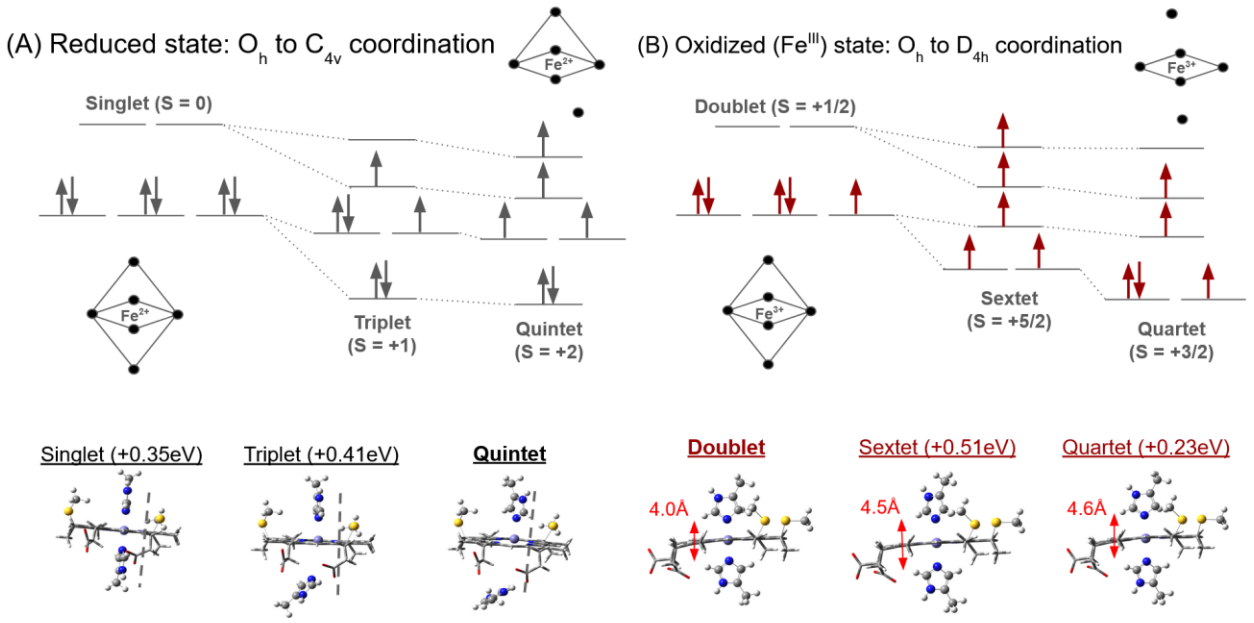


Figure 6.3 - The d-orbital splittings as predicted by crystal field theory as well as the relevant DFT optimized geometries, showing the transition to higher spin states when the octahedral symmetry is broken for (A) the reduced Fe^{II} state and (B) the oxidized Fe^{III} state. The bolded state represents the lowest energy spin configuration, splitting energies are given using this state as a reference.

Though the low spin state has the lowest electronic energy for each case, the higher spin states were found to be stable with only a small energy splitting (less than 1eV in all cases). While the geometry shift at higher spin states was more dramatic in the reduced case, with the coordinating histidine ligand shifting up to 4 Å away from the Fe center and rotating out of plane, the oxidized heme only experienced a small axial shift of the coordinating histidines (0.3 Å each) at higher spin states. The reorganization energies and redox potentials for these higher spin states are shown in Table 6, showing that these higher spin states have an increased energy barrier for charge hopping. The increase in reorganization energy by itself corresponds to a 10^5 fold decrease in the electron hopping rate. Of these higher spin states, the charge transfer between the triplet and quartet states

is the most favorable, with a comparable redox potential to the low spin state and slightly lower reorganization energy than the higher quintet to sextet transition. An analysis of the vibrational modes points towards the singlet to triplet transition having a much smaller crossover temperature (52 K) compared to all other transitions (calculated crossover temperatures exceeding 500 K). The spin transition temperature is calculated from the free energy of each spin state based on equation 46. While a calculation of the exact temperature of this transition will require benchmarking, the comparison shows that this spin transition would be most favorable for this proposed mechanism.

Table 6 - Reorganization energies and redox potentials for each of the possible spin state of electron transfer

Spin State Transition	**Reorganization Energy (eV)**	**Redox Potential (V)**
Singlet/Doublet (Low Spin)	0.093	-0.208
Triplet/Quartet	0.357	-0.373
Quintet/Sextet (High Spin)	0.353	-1.069

b. Limitations of collinear models

Many of the experimental setups used for verifying spin-filtering effects in cytochromes require addition of a magnetic field to create spin polarization of the current.[127,128,131] A common experimental setup for measuring spin-dependent transport is applying a magnetic field to a system containing ferromagnetic contacts, measuring the differences in current across varied magnetic

field directions. In the weak field limit, the corrective term to the Hamiltonian describing the Zeeman effect is given by the formula

$$\hat{H}_{zeeman} = -\hat{\vec{\mu}} \cdot \hat{\vec{B}} = \frac{\mu_B}{\hbar} \left(g_l \hat{\vec{L}} + g_s \hat{\vec{S}} \right) \cdot \hat{\vec{B}} \tag{40}$$

where $\hat{\vec{B}}$ is the magnetic field operator, $g_{\frac{l}{s}}$ are the gyromagnetic ratios for orbital/ spin angular momentum (respectively), μ_B is the Bohr magneton, $\hat{\vec{L}}$ is the orbital angular momentum operator, and $\hat{\vec{S}}$ is the spin angular momentum operator. The alignment of the angular momentum with the local magnetic field (given as a dot product) determines the strength of this interaction, which is not specified in a collinear system. Plugging values into equation 40, the induced Zeeman splitting of a d-orbital electron aligned with a 1 Tesla external field is 0.17 meV. The collinear model of electron spin is a non-relativistic approach that models spin interactions as purely exchange/correlational effects and excludes spin mixing terms. In this formulation, spin vectors are defined in a two-component basis and labeled as “alpha” and “beta” spins that have arbitrary directionality. Collinear spin transport is governed only by unpaired “alpha” spins, and spin direction is governed only by interaction with an external magnetic field. The result of this is that any spin dependence measured by changing the magnetic field cannot be explained by a purely collinear model: applying a magnetic field to the system will simply cause all alpha spin electrons to reorient. Differences in magnetic field strength can be accounted for through the Zeeman splitting term, but any change in direction of magnetic field with the cytochrome will cause realignment of the electron spins without any added energy penalty. Instead, spin dependence

caused by collinear models can be measured in experiments without an applied magnetic field, such as in the spin quantum Hall effect, or by controlling a spin-polarized current (for example, with a ferromagnetic contact) using a magnetic field independent of or isolated from the magnetic field applied to the cytochrome.

Further issues arise when using ab-initio methods like density functional theory (DFT) to model a characteristic spin Hamiltonian for the system. DFT is used to calculate the ground state of the system of interest, minimizing the energy of the system by populating orbitals from lowest to highest energy. In unrestricted DFT, alpha and beta electrons are treated with separated Kohn-Sham determinants, allowing their spatial distributions to differ. While this is a simple and straightforward description of open-shell systems, where the z-component of spin (S_z) remains well-defined, it can lead to cases where the electronic state is no longer a pure eigenstate of the total spin operator S^2. This is referred to as spin contamination and is typically measured by comparing the expectation value $\langle S^2 \rangle$ with the ideal value $S_z(S_z+1)$ for the target spin state. This was explored in detail in a Chapter 5, applying constraints to the spin contamination to minimize mixing of spin states in a restricted open-shell model.[129] Using this model, it was possible to consider spin states of a multi-heme cytochrome system to determine the ground state spin. While this approach is mathematically sound for a static system, its application to dynamic biological systems is a bit more complex. In such systems, coherence length scales can be much smaller than the system size, and a static ground state description may not accurately reflect the average

behavior of electron orbitals. Moreover, the restricted open-shell approach, while preventing spin contamination, can't properly describe certain important electronic configurations. These include broken-symmetry states, where different regions of a molecule have localized but antiferromagnetically coupled spins. While these broken-symmetry states aren't technically ground states (since they're not eigenstates of the total spin operator), they may potentially provide a better description of the real electronic structure in a system where coherent, delocalized states are not possible.[132] By definition these states are not ground states, because the spin operator does not commute with the Hamiltonian across the whole system, but they represent possible spin configurations that are impossible to achieve with restraints applied to the spin contamination.

c. Exploring non-collinear effects

To allow for spin mixing in the ab-initio calculations, constraints must be removed from the Fock operator. This is possible using the generalized open shell method, where the spin state of each orbital is fully unconstrained and can have arbitrary direction. That is, the $N \times N$ alpha and $N \times N$ beta matrix from an unrestricted open-shell calculation can be combined to a fully populated $2N \times 2N$ matrix containing onsite spin and spin mixing terms. Without adding relativistic effects, this treatment can be used to calculate spin frustration in a system (such as Cr and Fe clusters).[19] Instead of considering the fully relativistic Dirac equation with a four-component spinor, which is computationally intensive for molecular systems, expansions in the two-component spin basis can be used to model higher order spin effects.[134]

In a lightweight system such as a cytochrome, in the absence of an external magnetic or electric field, spin-orbit effects dominate the relativistic corrections. Still, the spin-orbit contribution from the Fe centers remains small, on the order of 10^{-5} eV for p-orbitals.[135] For comparison, the maximum dipole-dipole magnetic interaction between two electrons in separate hemes of a cytochrome is two orders of magnitude smaller, assuming a minimum Fe-Fe distance of 8 angstroms based on the crystal structure.[109] Even still, with the high density of states at the HOMO and LUMO bands, it is possible for orbital occupation to change significantly from slight orbital energy differences due to Coulomb and exchange interactions. Using the Breit Hamiltonian, which can be used as a correction to the non-relativistic Kohn-Sham equations, the Hamiltonian for spin-orbit can be separated into a single electron term, given by

$$H_{SO}^{(1)} = \frac{\alpha^2}{2} \sum_{i}^{elec} \sum_{n}^{atom} \frac{Z_j}{|\vec{r}_{in}|^3} (\vec{r}_{in} \times \overrightarrow{p_i}) \cdot \vec{S}_i \tag{41}$$

and a two-electron term given by

$$H_{SO}^{(2)} = \frac{\alpha^2}{2} \sum_{i}^{elec} \sum_{j>i}^{elec} \frac{1}{|\vec{r}_{ij}|^3} (\vec{r}_{ij} \times \overrightarrow{p_i}) \cdot (\vec{S}_i + 2\overrightarrow{S_j}) \tag{42}$$

where α is the fine structure constant, $\vec{r}_{in}$ is the direction vector for electron i relative to atom center n, $\vec{r}_{ij}$ is the direction vector for electron i relative to electron j, $\vec{p}_i$ is momentum of electron i, and $\vec{S}_i$ is the spin vector of electron i. From a computational perspective the two-center terms in equation 42 can be quite expensive, requiring two-electron integrals for each calculation. Instead,

this term can be approximated as a correction to the one-center term using integrals dependent on the angular momentum of each shell using the screen-nuclear-spin-orbit approximation. This method was applied to the Douglas-Kroll-Hess (DKH) expansion, which is a two-component expansion of the Dirac Hamiltonian rather than simply as a correction to the non-relativistic Hamiltonian, and has been shown to accurately predict spin-orbit splitting for f-block elements within 6% of full relativistic treatment.[134]

To look at non-collinear spin interactions between neighboring hemes, the previously modeled three heme subunit of the OmcS cytochrome was used,[129] which contains hemes in parallel and perpendicular conformations. Using this case as a reference, we analyzed the spin density across each heme center to look for non-collinear effects that may appear in longer multi-heme structures. This was done using the density matrix and finding the expectation values of each Pauli matrix to calculate the spin vector of each atomic orbital, given by

$$\langle \vec{S}_i \rangle = \left(\langle S_x \rangle_i, \langle S_y \rangle_i , \langle S_z \rangle_i \right) = \frac{\hbar}{2} \left(tr[\rho_i \sigma_x], tr[\rho_i \sigma_x], tr[\rho_i \sigma_x] \right) \tag{43}$$

where ρ_i is the 2x2 density matrix for an orbital given by

$$\rho_i = \begin{bmatrix} |\uparrow\rangle\langle\uparrow| & |\uparrow\rangle\langle\downarrow| \\ |\downarrow\rangle\langle\uparrow| & |\downarrow\rangle\langle\downarrow| \end{bmatrix} \tag{44}$$

The orbital spin vectors were summed up across each atomic basis function to generate local spin vectors for each atom. In addition to spin density, spin dependent transport was calculated using the formulation in equation 36. Recall that the addition of non-collinear effects and spin

mixing, the of-diagonal terms of $\hat{F}_{spin}$ are non-zero, leading to non-zero $T_{\uparrow\downarrow}$ and $T_{\downarrow\uparrow}$ transmissions that represent spin procession during transport across the cytochrome.

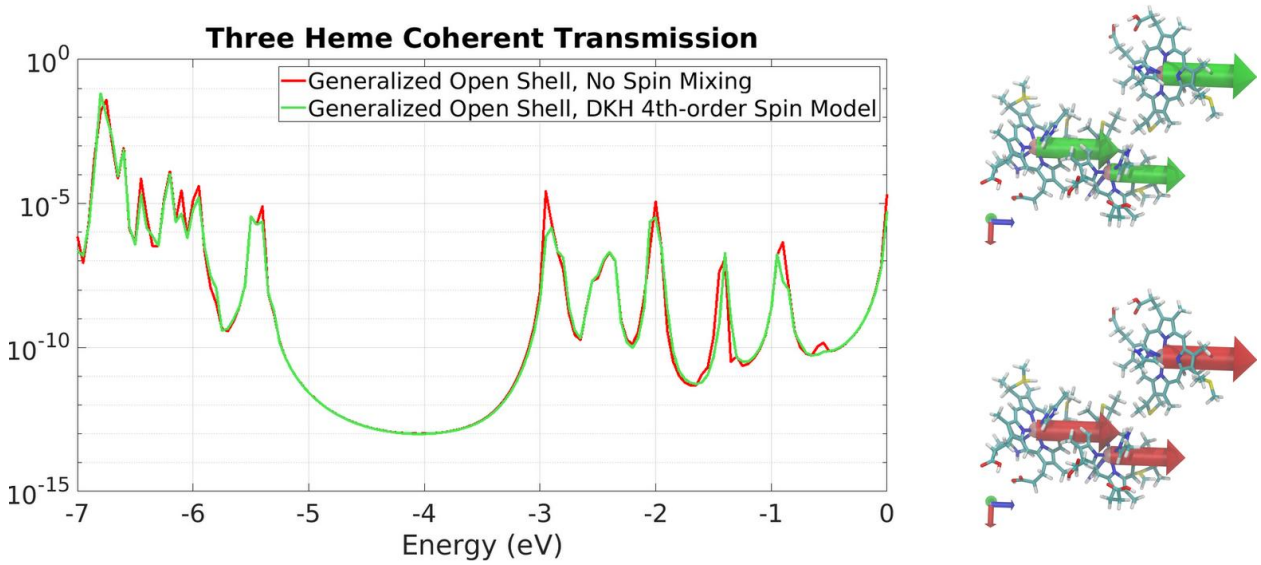


Figure 6.4 - Coherent transmission curves for the three-heme structure, calculated by adding constant self-energy contacts to coordinating histidine groups on the first and third heme, and corresponding atomic spin vectors, which only show up with significant magnitude on the Fe centers in the z-direction.

The spin vectors and transmission curves for the full calculation are shown in Figure 6.4, showing no noticeable change in the spin vector alignment (which are all aligned in the default z-direction) and only a very small change in the transmission curve. This shows that the addition of spin-orbit terms has a negligible effect on the alignment due to the light atoms present (Fe is the heaviest atom) and the lack of strong interactions between Fe sites, even in the parallel conformation case.

d. Modeling Methods

Ab-initio DFT calculations used the development version of the Gaussian DFT software package.[136] For all calculations on the single coordinated heme system, the B3LYP hybrid functional[47] was used with the mixed gaussian orbital basis set. This combined the LANL2DZ[60] basis set for Fe centers with the 6-31G** Pople basis[137] set for all other atoms (C, N, O, H, and S). All simulations were run using the polarizable continuum model as an implicit solvent, which models water as a continuous dielectric surface.[93] Total DFT electronic energies were used to calculate reorganization energy, redox potential, and spin state energy splittings. Open-shell DFT modeling was used for all states, with the unrestricted model used for collinear spin modeling and the generalized open-shell method used for non-collinear spin modeling.

The coordinated heme structure, shown in Figure 6.1C, was used for all optimization calculations by terminating the coordinating histidine and bonded cysteine groups at the beta carbon, as set up in previous works.[115,129] All heme centers were modeled with the propionate group in the deprotonated COO- state (-2 charge per heme), with the coordinate Fe group modeled in the reduced Fe^{II} (net zero charge) or oxidized Fe^{III} (net +1 charge) oxidation state. DFT optimization calculations did not apply any restraints to the coordinated heme structure and were found to be stable without the backbone structure. This coordinated heme structure was the same repeating unit used for the larger three-heme system in subsequent calculations.

For application of the non-collinear model to the three-heme system, the basis set was changed to the 6-31G** all electron basis for all sites. This was benchmarked against the mixed basis to ensure that the result did not change significantly and then used for generalized open shell calculations with and without the fourth order DKH expansion to model any spin-orbit effects. For each of these calculations, the ferromagnetic high spin state (quartet) was used as the collinear spin state constraint for the unrestricted case. This unrestricted case was used as an initial wavefunction guess for the generalized open-shell case, which was used as the initial wavefunction guess for the case with DKH spin-orbit terms.

Redox potentials were calculated using the ferro/ferricyanide system as a reference, which was chosen due to its chemical similarity to the coordinated heme system. Using the DFT setup mentioned previously, the change in DFT energy (3.13 eV) was added to the literature value of redox potential (0.37 eV)[138] to calculate the reference standard hydrogen electrode (SHE) energy (3.50 eV). Using this benchmark allowed us to cancel out any accumulated errors in the DFT energy that deviated from the literature SHE value (4.44 eV), adding the reference value to the DFT redox energy difference to calculate a redox potential.

To study the effects of geometry on the spin state in the nanowire, a single coordinated heme site was subjected to relaxation, varying the spin occupation. For this procedure the spin state was set as a boundary condition and the DFT energy was minimized by optimizing the geometry. For

the oxidized heme, a single heme in the doublet (spin = +1/2), quartet (spin = +3/2), and sextet (spin=+5/2) states were optimized and for the reduced case the single (spin = 0), triplet (spin = +1), and quintet (spin=+2) states were optimized. After optimization each structure was confirmed to be in the ground spin state by varying the spin state to confirm the lowest DFT energy occurred at the optimized state. The relative energies of each calculation are given in Table 7 below. Due to convergence issues only the unrestricted open-shell energy is reported, as secondary restricted open-shell calculations were not possible in all cases. Instead, spin contamination was monitored for each of the cases to ensure that all cases had a deviation of less than 0.001 from the expected S^2 value.

Table 7 - A comparison of the DFT energies for the oxidized and reduced case, all given relative to the lowest spin state and optimized geometry (bolded). The highlighted energies in the diagonal are the energies of the electron state used to optimize the geometry structure given by the first row. Energies in black are the Q=-2 charge state and energies in red are the Q=-1 state, each column has a different spin multiplicity.

Optimization Geometry	**Relative Energy (eV) for each Charge/Spin State**					
	Singlet (Q=-2, M=1)	Triplet (Q=-2, M=3)	Quintet (Q=-2, M=5)	Doublet (Q=-1, M=2)	Quartet (Q=-1, M=4)	Sextet (Q=-1, M=6)
Reduced Singlet	0.35	1.22	1.32	0.10		
Reduced Triplet	1.06	0.41	0.68		0.55	
Reduced Quinet	0.79	0.29	**0.00**			0.46
Oxidized Doublet	0.44			**0.00**	1.89	2.64
Oxidized Quartet		0.77		0.69	0.23	0.67
Oxidized Sextet			0.76	0.80	0.34	0.51

After optimization, a vibration calculation was conducted to ensure structure stability and generate thermochemistry data. The geometries were also compared between each structure, looking at the point group and symmetry of the coordinated Fe (such as the N-Fe bond lengths) to explain the results. Using these structures, the reorganization energies were calculated using the formula

$$\lambda = \frac{1}{2}([E^A_{red} - E^D_{red}] + [E^D_{ox} - E^A_{ox}]) \quad (45)$$

where D and A refer to the donor and acceptor minimized geometries respectively and subscripts red and ox refer to the reduced and oxidized electronic states respectively. The entropy data from the vibrational calculations were used to calculate the spin transition temperatures for each pair of states. The spin transition temperature, T_{spin} is defined as the temperature where the two states are in equilibrium, as given by

$$G_1 = H_1 + T_{spin}S_1 = H_2 + T_{spin}S_2 = G_2 \quad (46)$$

where, for spin state i, G_i is the free energy, H_i is the enthalpy, and S_i is the entropy are thermochemistry properties calculated in the Gaussian package.[139]

e. Discussion

We presented some new insights into cytochrome spin transport, first looking in detail at reorganization in a single heme. The energy barriers that correspond to a spin transition at higher

energy are large and could explain large spin filter effects exhibited in experimental results. Further benchmarking could be helpful, especially comparing experimental absorption spectra to ab-initio results. The vibrational modes associated with symmetry breaking in the heme system correspond to frequencies in the range of 160-400 cm^{-1}, with prominent peaks shown in Figure 6.5 along with the vibrational mode for each spin state. Comparing these vibrational peaks, which are dependent on the distortions of coordinating histidines in the different spin states, may help identify the spin state of the material and understand the reorganization effects in the system.

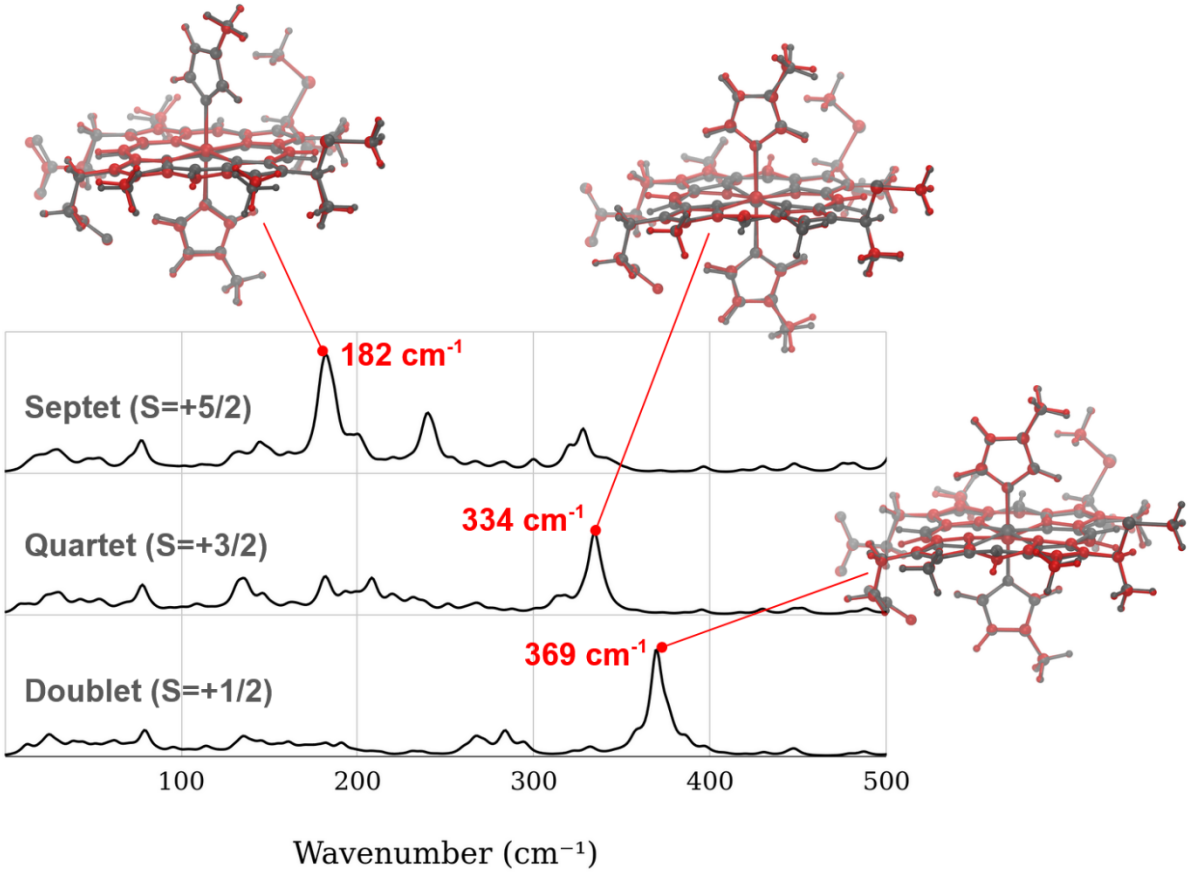


Figure 6.5 - Generated IR absorption spectra for the oxidized cases showing the axial histidine vibrational peaks for each of the three modeled spin states. Inset structures show structure stretching across the vibration mode corresponding to the peak.

Regardless of the ab-initio calculation setup, the connection with magnetic field dependence is still unclear from all modeling results. The addition of spin-orbit effects did not shift any of the transmission peaks or atomic spin vectors noticeably, even including the full fourth-order DKH model. As an alternative explanation, we can look at the interaction between the Au surface and the nanowire during magnetic field-dependent measurements. As a much heavier atom the intrinsic spin-orbit interaction is much stronger and can lead to significant spin polarization of surface

states.[140] We suspect that the interaction of this orbital with the heme sites on the OmcS nanowire aligns the unpaired spins with the plane of the Au surface, creating a spin dependence that is resilient to the effects of a weak magnetic field. This realignment would result in measured current that is less dependent on the magnetic field, rather than the fully paired d-orbitals of the Au atoms that would not be impacted by the changing magnetic field. Exploring this contact interface will help understand the experimental results, especially to distinguish between effects intrinsic to the cytochromes or induced by the bonded metal atoms.

7. Contact-Induced Spin Selectivity in Cytochromes

The quantum mechanical property of spin, long studied in solid-state physics, has emerged as a crucial factor in biological systems, revealing unexpected connections between quantum phenomena and life processes. Studies have demonstrated that cellular development responds to the earth's magnetic field,[141] while critical biological processes like cryptochrome reactions exhibit clear spin-dependent behavior.[131] This quantum-biological interface extends to the molecular level, where chiral structures can act as spin filters[142] – a phenomenon particularly well-documented in double-stranded DNA.[143,144] This property, called chiral induced spin selectivity (CISS), has been described as an intrinsic property of the materials, where the momentum of a helical electron path couples with spin.[145] This phenomenon has been attributed to spin-orbit coupling, creating spin-mixing terms that induce spin current, but the magnitude of the effect suggests that interactions

with contacts have high contributions to the spin filtering effect.[56] In addition, recent experimental results show that cytochromes without significant chirality exhibit significant magnetoresistance,[128] showing that non-collinear spin interactions with electrical contacts may significantly contribute to spin filtering effects, an effect that has been demonstrated theoretically in small organic systems.[146] These models use a non-equilibrium Green's function (NEGF) approach to model electrical contacts in a nanodevice, using the spin dependent quantum transmission to calculate spin currents through a molecule.

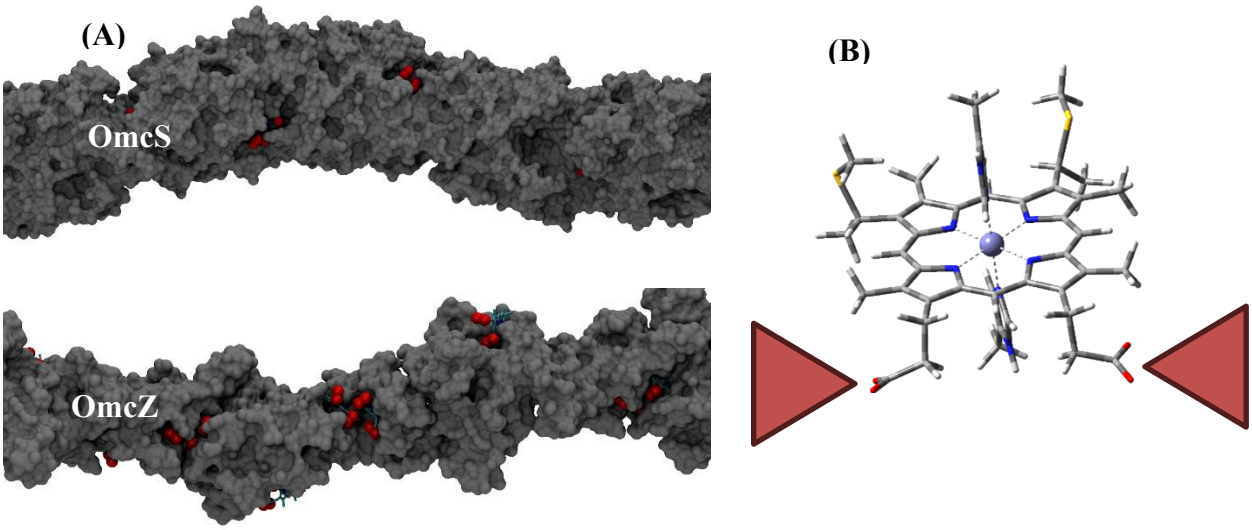


Figure 7.1 – (A) The location of the propionate groups (carboxylate oxygens shown in red) in the OmcS (PDB ID 6EF8)[109] and OmcZ (PDB ID 7LQ5)[147] cytochromes. (B) The coordinated heme system model with two contacts (one on each propionate group).

Building on our previous work in Chapter 5, where a collinear spin transport model was applied to cytochromes, we have extended our model to include non-collinear effects caused by contact interactions. In the outer membrane cytochromes, electron transport occurs through coordinated heme groups. Based on the geometries of the hemes in such systems, as depicted in Figure 7.1A,

it is likely that the propionate groups on the heme-B molecule would most likely interact with electrodes. Therefore, our model will be applied to contacts interacting with the propionate moiety on the heme. The simplified extended system will use a single heme with Au contacts, as shown in Figure 7.1B, to bridge theory with emerging experimental results conducted on Au substrates.[127,128]

For this investigation, we have developed a software package using self-consistent NEGF density function theory (NEGF-DFT) with the Gaussian software package to model spin transport in biological systems. This work uses many algorithms from the previously developed ANT.Gaussian software package and one algorithm from reference [13] built with Gaussian 03 and 09. The main differences are: (1), porting this code to use the *gauopen* python interface developed with the newest Gaussian releases and (2), adding generalized open-shell compatibility (with complex coefficients). To understand the effects of contact type and contact geometry, we first investigated the phenyl dithiol (PDT) molecule. This was chosen due to the large amount of modeling[148–151] and experimental[152] work done with this system. After validating the model with the PDT system, the software package was used to model a single heme system, specifically focusing on spin-orbit effects between the contact and the molecule.

a. Theory

The NEGF-DFT software package, GauNEGF, was built using the Numpy[153] and Scipy[154] Python packages to conduct all NEGF calculations and the Gaussian development version to run

all DFT calculations, which include spin-orbit and solvent effects. The software package has been published on Zenodo[155] and includes an API and examples for reference.

Non-equilibrium Green's Function (NEGF) Formalism

The NEGF formalism was applied to the given molecule with two electrical contacts modeled as a two-terminal device, with the retarded Green's function $\hat{G}^R$ defined previously in equation 27, and the advanced Greens function given by

$$\hat{G}^{\mathrm{A}} = \left(\hat{G}^R\right)^\dagger \tag{47}$$

which can be written as a matrix in the specified basis. The retarded Greens function is represented as a matrix in the atomic orbital basis, which can be calculated at any energy E. Similarly, self-energies are matrices representing the electrical contacts given by the formula

$$\hat{\Sigma}_i^r = \hat{\tau}_i \hat{\mathrm{g}}_{s,i}^r \hat{\tau}_i{}^\dagger \tag{48}$$

where τ_i is the Hamiltonian coupling matrices to the electrical contact i and $\hat{\mathrm{g}}_{s,i}^r$ is the surface green's function of the contact i. To calculate the coherent electron transmission though the molecular system, we can apply equation 30 to a two-contact system, given by

$$T(E) = tr\left[\hat{\Gamma}_L \hat{G}^r \hat{\Gamma}_R \hat{G}^a\right] \tag{49}$$

where $\hat{\Gamma}_{L/R}$ are the contact couplings given by

$$\hat{\Gamma}_i = i\left(\hat{\Sigma}_i^r - \hat{\Sigma}_i^a\right) = 2Im\left[\hat{\Sigma}_i^r\right] \tag{50}$$

The transmission function can be used to calculate the electrical current in the device which is given by the Landauer formula

$$I(\mu_L, \mu_R) = -\frac{q_e^2}{h} \int_{-\infty}^{\infty} T(E)\,[f_L(E, \mu_L) - f_R(E, \mu_R)]\,dE \tag{51}$$

for a voltage $\Delta V = \frac{(\mu_L - \mu_R)}{q_e}$ applied, with q_e defined as the magnitude of the charge of an electron, and with Fermi functions $f_i(E, \mu_i)$ given by

$$f_i(E, \mu_i) = \frac{1}{\exp\left[\frac{E - \mu_i}{k_B T}\right] + 1} \tag{52}$$

where k_B is the Boltzmann constant. In previous chapters the Hamiltonian of the system was generated from the single-electron Fock and overlap matrices, which can be plugged into equation 49 to calculate the transmission. While this approximation works well for a molecular system with weakly coupled contacts, using the Hamiltonian of the isolated molecule does not consider the effect of the contacts on the molecular orbitals. This becomes especially important under an applied bias, where onsite orbital energies can shift to match the Fermi energy of the contacts. Furthermore, the current is a function of the Fermi energy of the system, which is highly dependent on molecule contact interactions that are not included in the isolated system calculations.

NEGF-DFT Implementation

To fully include the effects of the contacts on the molecular device, the NEGF formalism can be applied to generate the density matrix self-consistently with DFT, called NEGF-DFT. In the

standard Kohn–Sham DFT implementation, a single self-consistent field (SCF) cycle involves solving the Kohn-Sham equations using an input density matrix to generate a single-electron Fock matrix, which can be diagonalized to generate Kohn-Sham orbitals.[48] The generated orbitals are then populated from lowest to highest energy to generate the ground state of the multi-electron system, which can be represented by a density matrix. The occupied orbitals can each have a pair of electrons without a spin index (called the restricted close shell case) or can have individual electrons with spin (called the open-shell case) with interactions parameterized by the exchange-correlation functional. In the case of a molecule coupled with electrical contacts, rather than populating the orbitals in the ground state of the isolated system, the density matrix can be constructed using the non-equilibrium Green's function of the full system (device and contacts) as shown in Figure 7.2.

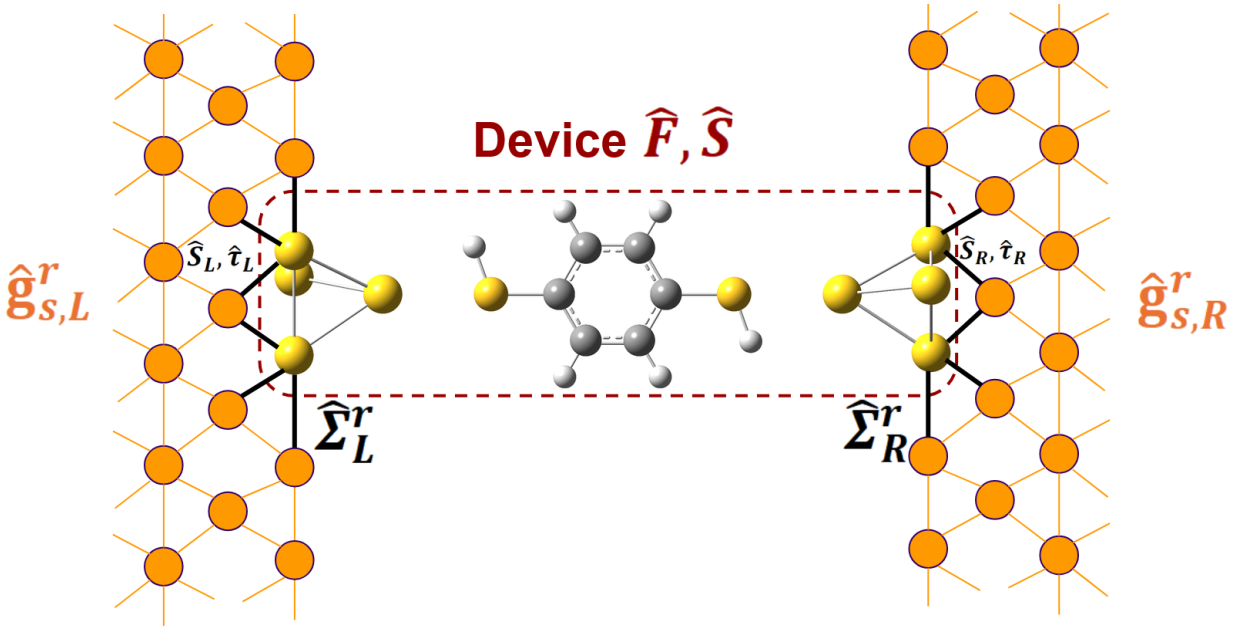


Figure 7.2 - Device setup for Green's function and self-energy calculations, showing the extended system with Au contacts and surface Green's function of the full semi-infinite Au surface.

Using the single electron Fock matrix from DFT, $\hat{F}$, we can construct the retarded Green's function

$$\hat{G}^r = \left(E\hat{S} - \hat{F} - \hat{\Sigma}^r_L - \hat{\Sigma}^r_R\right)^{-1} = \left(\hat{G}^a\right)^\dagger \tag{53}$$

where $\hat{S}$ is the overlap matrix of the atomic orbital basis.[149,156] Similarly, we can construct the self-energy matrices to be

$$\hat{\Sigma}^r_i = \left(E\hat{S}_i - \hat{\tau}_i\right)\hat{g}^r_{s,i}\left(E\hat{S}_i - \hat{\tau}_i\right)^\dagger \tag{54}$$

where $\hat{\tau}_i$ and $\hat{S}_i$ are the Fock and Overlap sub-matrices that couple contact i and the device (see Figure 7.2). Note that for a fully orthogonal basis set, the $\hat{S}_i$ used in the equation above becomes

zero, transforming it into equation 48, while $\hat{S}$ becomes the identity matrix. Plugging this into equation 50 to get $\hat{\Gamma}_i$ we can construct the less-than Green's function for the device given by

$$-i\hat{G}^{<} = \hat{G}^{r}\left(f_L(E,\mu_L)\hat{\Gamma}_L + f_R(E,\mu_R)\hat{\Gamma}_R\right)\hat{G}^{a} \tag{55}$$

Integrating $-i\hat{G}^{<}$ as a spectral function, we can calculate the density matrix $\hat{\rho}$ to be

$$\hat{\rho} = \frac{1}{2\pi}\int_{-\infty}^{\infty} -i\hat{G}^{<}\,dE \tag{56}$$

This density matrix is constructed at each cycle as the input to the density functional to self-consistently solve for the device Hamiltonian. Convergence of the NEGF-DFT equations occurs when the criteria

$$\hat{\rho}\left[\hat{F}[\hat{\rho}]\right] = \hat{\rho} \tag{57}$$

is met, meaning that the Fock matrix generated from the density functional can be plugged into equation 56 to generate a matching density matrix. While this formalism can accurately describe a nanodevice with electrical contacts, computational intensity increases dramatically for the integration in equation 56 especially for the cases of energy-dependent contacts.

Density Matrix Calculation

There are two key cases that can reduce the computational cost of density integration: (a) non-equilibrium, energy-independent self-energy and (b) equilibrium, energy-dependent self-energy. In the case of energy-independent self-energy, the integral in equation simplifies to an analytical solution at the low temperature limit, only requiring matrix computations at the limits of

integration as realized first by Damle [13]. This case has been implemented previously for both for PDT and for a gold nanowire.[149] In the case of equilibrium (zero bias applied, $\mu_L = \mu_R$), the integral also simplifies significantly and can be computed efficiently in the complex plane. This method has been used in NEGF-DFT packages,[156–158] separating the density matrix into an equilibrium and non-equilibrium contribution.

Energy Independent Self-energy

Following the method described previously by Damle et al,[159] we will start by separating the density contribution from each contact ($j \in L, R$) to get

$$2\pi\hat{\rho}_j = \int_{-\infty}^{\infty} \hat{G}^R \, \hat{\Gamma}_j \hat{G}^A f_j(E, \mu_j) dE \tag{58}$$

$$\hat{\rho} = \hat{\rho}_L + \hat{\rho}_R$$

Applying an inverse Löwdin transformation to orthonormalize the Green's function operator we get

$$2\pi\rho_j = \hat{S}^{-\frac{1}{2}} \left[\int_{-\infty}^{\infty} \bar{G}\bar{\Gamma}_j \bar{G}^\dagger f_j(E, \mu_j) dE \right] \hat{S}^{-\frac{1}{2}} \tag{59}$$

with transformed variables given by

$$\bar{\Gamma}_j = \hat{S}^{-\frac{1}{2}} \hat{\Gamma}_j \hat{S}^{-\frac{1}{2}} \tag{60}$$

$$\bar{G} = \left(E\hat{I} - \bar{F}\right)^{-1} \tag{61}$$

$$\bar{\mathrm{F}} = \hat{S}^{-\frac{1}{2}}\left(\hat{F} + \hat{\Sigma}_L + \hat{\Sigma}_R\right)\hat{S}^{-\frac{1}{2}} \tag{62}$$

In this orthonormal basis we can rewrite the transformed Green's function in terms of the eigenvectors of $\bar{\mathrm{F}}$

$$\bar{\mathrm{F}}\,|n\rangle = \epsilon_n|n\rangle \tag{63}$$

to get

$$2\pi\rho_j = \hat{S}^{-\frac{1}{2}}\left[\int_{-\infty}^{\infty} f_j(E,\mu_j)\left(\sum_n \frac{|n\rangle}{E-\epsilon_n}\right)\langle n|\bar{\Gamma}_{\mathrm{j}}|n\rangle\left(\sum_{n'}\frac{\langle n'|}{E-\epsilon_n^{\dagger}}\right)dE\right]\hat{S}^{-\frac{1}{2}} \tag{64}$$

Using this form, several simplifications can be made. At zero temperature, we can convert the integral to have finite limits,

$$2\pi\rho_j = \hat{S}^{-\frac{1}{2}}\left[\int_{E_{lower}}^{\mu_j}\left(\sum_n \frac{|n\rangle}{E-\epsilon_n}\right)\langle n|\bar{\Gamma}_{\mathrm{j}}|n'\rangle\left(\sum_{n'}\frac{\langle n'|}{E-\epsilon_n^{\dagger}}\right)dE\right]\hat{S}^{-\frac{1}{2}} \tag{65}$$

where E_{lower} represents a lower energy cutoff for integration. In addition, if we assume that the self-energy of the contact is energy-independent, the integral can be solved analytically for each matrix element within the $\bar{\mathrm{F}}$ basis:

$$\langle n|2\pi\rho_j|n'\rangle = \frac{\langle n|\bar{\Gamma}_{\mathrm{j}}|n'\rangle}{\epsilon_n - \epsilon_{n'}^{\dagger}}\left[\log\left(\frac{\mu_j-\epsilon_n}{\mu_j-\epsilon_{n'}^{\dagger}}\right) - \log\left(\frac{E_{min}-\epsilon_n}{E_{min}-\epsilon_{n'}^{\dagger}}\right)\right] \tag{66}$$

This expression can be applied in both equilibrium and non-equilibrium cases, summing up the contributions from each contact at different Fermi energies.

Energy Dependent Self-energy

Using the full form of the self-energy matrix from equation 48 requires numerical methods, as the integral cannot be simplified analytically. However, the equilibrium case, that is $\mu_R = \mu_L$, allows for simplification of the integration step in equation 56, where $\hat{G}^<$ can be reduced to

$$i\hat{G}_{eq}^{<} = i\left(\hat{G}^R - \hat{G}^A\right) = -2Im\left(\hat{G}^R\right) \tag{67}$$

based on the properties of the spectral function. Substituting this into equation 56 after noting that the Fermi function is real gives

$$\hat{\rho}_{eq}(\mu) = -\frac{1}{\pi} Im\left[\int_{-\infty}^{\infty} \hat{G}^R f(E,\mu)\, dE\right] \tag{68}$$

Since the retarded Green's function only has poles in the lower half of the complex plane and we restrict ourselves to zero temperature when the Fermi function has no poles, the integrand in the above equation is well behaved in the upper half of the complex plane. We numerically integrate the above with a contour in the upper complex plane by applying Cauchy's theorem,

$$\hat{\rho}_{eq}(\mu) = -\frac{1}{\pi} Im\left[\int_{-\infty}^{\mu} \hat{G}^R\, dE\right] \tag{69}$$

Fermi Energy Calculation

Calculation of the Fermi energy is consistently a challenge for NEGF calculations. In this work, we have calculated the Fermi energy for the contact assuming charge neutrality on the

contact atoms and allowing charge to flow from the contacts to the device. This Fermi energy is calculated independent of electron spin, as all contacts used for this study are spin independent. To impose charge neutrality, we use the following boundary condition:

$$Tr\left[\hat{\rho}_{eq}(\mu)\hat{S}\right] - N_{elec} = N(\mu) - N_{elec} = 0 \tag{70}$$

at equilibrium. For the energy-independent case, calculation of the density matrix as a function of Fermi energy is a simple task, only requiring a single diagonalization step and a single inversion step for all calculations, as the Fermi energy only impacts the multiplication constant in equation 66. Therefore, the Fermi energy can be calculated simply using the bisection method, converging in a finite number of iterations independent of the DOS profile.

A similar approach was developed for the energy-dependent case, using a constant self-energy approximation to predict the shift in Fermi energy required to create electroneutrality (equation 70). To use this approximation, we first calculate the total number of electrons given this constant self-energy,

$$N_{const}(\mu_{\mathrm{i}}) = Tr\left[\hat{\rho}_{\mathrm{const}}(\mu_{\mathrm{i}}) \times \hat{S}\right] \tag{71}$$

where μ_{I} is the current value of the Fermi energy, and then we solve

$$N_{const}(\mu_{i+1}) - N_{const}(\mu_i) = N_{elec} - N(\mu_i) \tag{72}$$

for the value of $\mu_{\mathrm{i+1}}$, the new Fermi energy prediction. The terms with the subscript 'constant' indicate that they are calculated with a constant self-energy for both contacts. Note that the RHS of equation (26), $N(\mu_i)$ refers to the number of electrons calculated with the energy-dependent self-energy. While this first order approximation does not give an exact solution at each iteration,

and requires extra iterations for convergence, the value of the Fermi energy does converge and can be a significant time savings over direct numerical integration. To calculate $N_{const}(\mu_i)$, we use the semi-analytical approach developed by Damle (see equation 66). This approach worked well when the Fermi energy did not lie on a peak at the density of states. When the Fermi energy lies on a peak in DOS, or similar instabilities exist in the system, both the secant method and Muller method were implemented for root finding in equation 70, to benchmark the constant self-energy method. These implementations were built to match the algorithm in the ANT.Gaussian code exactly,[160] updating the contact Fermi energy at each iteration before density matrix calculation.

In a previous attempt to calculate the Fermi energy without full numerical integration, a Taylor expansion of the DOS was used to create an nth order approximation of the function given by

$$N(\mu) = N(E) + N'(E)(E-\mu) + \cdots + \frac{N^{(n)}(E)(E-\mu)^n}{n!} = N_{elec} \tag{73}$$

Using the relation

$$N'(E) = DOS(E) = \frac{Tr\left[Im\left[\hat{G}^R(E)\right]\right]}{\pi} \tag{74}$$

calculating these coefficients could be done using n energy points to solve the set of linear equations given by

$$DOS(E+\Delta E_i) = DOS(E) + DOS'(E)(\Delta E_i) + \cdots + \frac{DOS^{(n)}(E)(\Delta E_i)^n}{n!} \tag{75}$$

This method was implemented in the `DOSFermiSearch.py` package, applying this search algorithm at each cycle of the NEGF-DFT calculation and adjusting the Fermi energy by the nth order prediction until convergence is reached. While this method was mathematically sound, eventually converging to the system Fermi energy for some stable cases, high non-linearity of the DOS profile made this method unfeasible for most molecules. Additionally, this method required n matrix inversion steps at each cycle compared to the single diagonalization and inversion step required for the constant self-energy approximation.

Combined Algorithm

In our implementation, the density matrix was computed through several integration steps. For the energy-independent case, calculations can simply be conducted by two analytical integrals, given by equation 66, from a minimum cutoff set to a very low energy (E_{lower} in Figure 7.3, set to -10^5 eV). For full energy-dependent contacts, a three-step integration procedure is required. First, a cutoff is calculated below the lowest energy eigenvalue, labeled as E_{min}. This is set as the upper limit for the first integration step, shown as contour A in Figure 7.3, which occurs on the real axis using the equilibrium density matrix

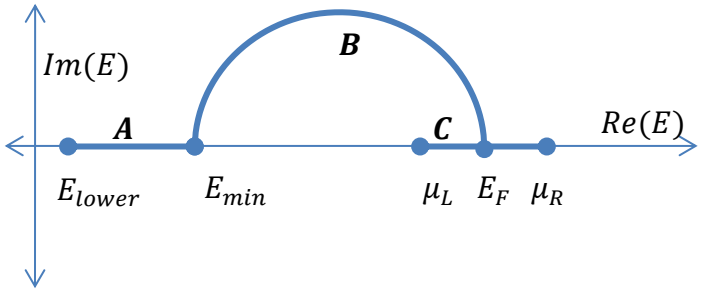


Figure 7.3 - The three-step integration contour used for density matrix calculation. Note that at equilibrium $\mu_L = \mu_R = E_F$ and contour C is not used.

equation given by equation 58. Note that because there are no poles between E_{lower} and E_{min}, it is more efficient to integrate on the real axis. Then, a semicircular complex contour is used to integrate the same equation from E_{min} to the equilibrium Fermi energy, E_F, shown as contour B in Figure 7.3. For non-equilibrium cases, that is $\mu_R \neq \mu_L$, the third step uses two non-equilibrium integrals on the real axis from E_F to μ_L and from E_F to μ_R, shown as contour C in Figure 7.3, using equation 69 to get

$$-\frac{1}{\pi}\int_{E_{lower}}^{E_{min}} Im[G^r(E)]dE - \frac{1}{\pi}\, Im\left[\int_{E_{min}}^{E_F} G^r(E_z)dE_z\right] - \frac{1}{2\pi}\int_{E_F}^{\mu_L} \hat{G}^R\, \hat{\Gamma}_R \hat{G}^A dE + \frac{1}{2\pi}\int_{E_F}^{\mu_R} \hat{G}^R\, \hat{\Gamma}_R \hat{G}^A dE \tag{76}$$

where the 2nd integral is calculated over the complex plane using contour B, while the other three integrals are evaluated over the real axis. E_z respresents the complex energy. Each definite integral was computed numerically using a weighted quadrature, with the formula given by

$$\hat{\rho}_i = \int_a^b d\hat{\rho}_i(z)dz \approx \sum_{j=1}^{n} w_j\, d\hat{\rho}_i\left(\frac{b-a}{2}x_j + \frac{a+b}{2}\right)\left(\frac{dz}{dx_j}\right) = \bar{\rho}_i(n) \tag{77}$$

where w_j are the weights applied to the points x_j in the n[th]-order quadrature, and $\bar{\rho}_i(n)$ is the contribution to the density matrix from the approximated definite integral on contour $i \in \{A, B, C\}$. For the real axis integrals, a Gauss-Legendre quadrature was used, while for the semi-circular complex contour a modified Gauss-Chebyshev quadrature was used based on the implementation in ANT.Gaussian.[160] The points and weights used for various quadrature types are given in the following table:

Type of Quadrature	x_j **definition**	w_j **definition**	***gauNEGF* usage**
Reimann Sum (midpoint rule)	$\frac{2n}{j} - 1$	$\frac{2}{n}$	Not used by default
Gauss-Legendre	j[th] root of the Legendre polynomial $P_n(x)$	$\frac{2}{(1-x_j^2)[P_n'(x_j)]^2}$	Used for all density matrix integration along the real axis
Gauss-Chebyshev	$\cos\left(\frac{j\pi}{n+1}\right)$	$\frac{\pi}{n+1}\sin^2\left(\frac{j\pi}{n+1}\right)\frac{1}{\sqrt{1-x^2}}$	Not used by default
ANT.Gaussian (modified Gauss-Chebyshev)	$\theta_j = \frac{(2j+1)\pi}{2n}$ $x_j = 1 + 0.21\sin(\theta_j)\cos(\theta_j) \times (3 + 2\sin^2(\theta_j)) - \frac{2j+1}{n}$	$\frac{16\sin^4(\theta_j)}{3n}$	Used for semicircular complex contour integration

The contours and corresponding quadratures were chosen to reduce the number of grid points required for integration of the density matrix while maintaining accuracy. The grid density is calculated for each system for a given tolerance δN such that

$$\max_j \langle\psi_j|\bar{\rho}_i(2n) - \bar{\rho}_i(n)|\psi_j\rangle < \delta N \tag{78}$$

for all contours. The value of n is used to determine the order of the quadrature in equation 77 to use for numerical integration. Similarly, the lower limit of integration for the numerical integration, E_{min}, is chosen such that

$$\max_j \langle \psi_j | d\hat{\rho}_i(E_{min}) | \psi_j \rangle < \delta N \tag{79}$$

This search starts at the lowest orbital energy and decreases to lower energies to ensure that contour A remains sufficiently far from complex poles of $\hat{G}^R(z)$. By default, the δN cutoff is set to 1E-4 and the starting guess for E_{min} is set to 10 eV below the lowest occupied orbital energy. This E_{min} typically passes the inequality in equation 79 without lowering the energy further and typically requires an integration grid with about $n = 128$ points to pass the inequality in equation 78.

The integration path and limits were chosen to reduce the energy grid requirements required for integration by avoiding poles, especially those of $\hat{G}^r$ found in the lower half of the complex plane. This becomes more of a challenge when accounting for non-zero temperature due to the addition of poles from the Fermi function (which do not exist for the T=0 definite integral limit). We can look at the Fermi function formula in equation 52 to see that there exist infinite poles given by

$$z_{\mu_i} = \mu_i + \mathrm{i}\,\pi k_B T(2n+1) \text{ for } n \in \mathbb{Z}$$

for both the left and right contacts. Therefore, to account for temperature, integration over a larger energy window along the real axis was required for contour B and C. To account for the tail of the Fermi function, each integral is extended by $5k_BT$ along the real axis and the contacts are

sorted into contact H, the contact with the lower non-equilibrium Fermi energy, and contact L, the contact with the higher non-equilibrium Fermi energy. This modified equation 76 to become

$$-\frac{1}{\pi}\int_{E_{lower}}^{E_{min}} Im[G^r(E)]dE - \frac{1}{\pi}\, Im\left[\int_{E_{min}}^{E_F} G^r(E_z)f(E,E_F)dE_z\right] - \frac{1}{\pi}\int_{E_F}^{E_F+5k_BT} Im[G^r(E)]dE$$
$$-\frac{1}{2\pi}\int_{\mu_L-5k_BT}^{E_F+5k_BT} \hat{G}^R\,\hat{\Gamma}_H\hat{G}^A[f(E,E_F)-f(E,\mu_L)]dE$$
$$+\frac{1}{2\pi}\int_{E_F-5k_BT}^{\mu_H+5k_BT} \hat{G}^R\,\hat{\Gamma}_H\hat{G}^A[f(E,\mu_H)-f(E,E_F)]dE \tag{80}$$

To ensure stability, the density matrix at each iteration, ρ_i, was incrementally formed at each iteration with a linear combination of the density matrix calculated at the current iteration (the solution of the integral in equation 85 using the previous Fock matrix extracted from Gaussian, given by ρ_{calc}) and the matrix at the previous iteration, ρ_{i-1}, using the formula

$$\rho_i = \alpha\rho_{calc} + (1-\alpha)\rho_{i-1} \tag{81}$$

where α is the mixing factor (default value of 0.02). Additionally, a Pulay mixing scheme was applied every $n+1$ iterations using calculated density matrices for the previous n guesses (by default $n = 4$). To calculate these coefficients, the inner product of the change in density matrix was calculated, given by

$$B_{ij} = \langle \rho_i - \rho_{i-1}, \rho_j - \rho_{j-1} \rangle \tag{82}$$

where B_{ij} is a single complex value given by summing all element-wise density matrix variations between iterations i and j. These values are used to solve the matrix equation

$$\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} & -1 \\ B_{12} & B_{22} & \cdots & B_{2n} & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{1n} & B_{1n} & \cdots & B_{nn} & -1 \\ -1 & -1 & \cdots & -1 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix} \tag{83}$$

for the Pulay coefficients c_n to update the density matrix to

$$\rho_{Pulay} = \sum_{i=1}^{n} c_i \rho_i \tag{84}$$

as defined in the original work.[161] Creating a Pulay "kick" step every n iterations matches the implementation in ANT.Gaussian,[156] with the added feature of allowing for complex Pulay coefficients in open-shell cases (where the density matrix can have a non-zero imaginary component).

Bethe Lattice Contact Model

The Bethe lattice contact was implemented based on the implementation in ANT.Gaussian.[160] Contacts were modeled as the (111) surface of an FCC lattice, using tight binding parameters for the 12 nearest neighbors to generate the self-energy of the bulk material. Using the Bethe lattice approximation, the bulk self-energy is given by the series of coupled equations

$$\Sigma_{\mathrm{k}}(E) = V_k \left(E - H - \sum_{j \neq k} \Sigma_j(E) \right)^{-1} V_k \tag{85}$$

where k is defined for each of the 12 nearest neighbors with coupling matrix V_k, the onsite Hamiltonian of a single atom is given by H, and E is the energy value. In this implementation, the

onsite Hamiltonian and coupling matrices are defined by orthonormal Slater-Koster parameters.[162] In this work we solved the self-energy at the surface considering the 6 in-plane nearest neighbors on the surface, each defined by a surface self-energy Σ_l^{surf}, and 3 out of plane neighbors, each defined by a bulk self-energy Σ_k, given by the equations

$$\Sigma_l^{surf}(E) = V_k\left(E - H - \sum_{m\neq l}\Sigma_l^{surf}(E) - \sum_k \Sigma_k(E)\right)^{-1} V_k \tag{86}$$

using the same variables given in equation 85. Contacts are constructed summing each contribution to a given contact atom, with Σ_l^{surf} generated for each nearest neighbor atom on the surface plane and Σ_k added for each of the 3 out of plane nearest neighbors. An example of such a contact is shown in Figure 7.4, which is a four-atom system with a single Bethe lattice contact applied to three of the atoms. For this system, the self-energy matrix corresponding to the single contact is given by

$$\Sigma_{bethe} = \begin{bmatrix} \Sigma_1^{surf} + \Sigma_2^{surf} + \Sigma_5^{surf} & 0 & 0 & 0 \\ +\Sigma_6^{surf} + \Sigma_1 + \Sigma_2 + \Sigma_3 & & & \\ 0 & \Sigma_1^{surf} + \Sigma_2^{surf} + \Sigma_3^{surf} & 0 & 0 \\ & +\Sigma_4^{surf} + \Sigma_1 + \Sigma_2 + \Sigma_3 & & \\ 0 & 0 & \Sigma_3^{surf} + \Sigma_4^{surf} + \Sigma_5^{surf} & 0 \\ & & +\Sigma_6^{surf} + \Sigma_1 + \Sigma_2 + \Sigma_3 & \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{87}$$

where each surface self-energy matrix is calculated from equation 86 and each bulk self-energy is calculated from equation 85. To use the basis with the non-orthogonal Gaussian orbital basis for

Green's function calculations, the Bethe contact is deorthagonalized using a Löwdin transformation given by

$$\Sigma_{contact} = S^{\frac{1}{2}}(\Sigma_{bethe})S^{\frac{1}{2}} \tag{88}$$

where S is the overlap matrix of the device (see Figure 7.2) as suggested by reference 27 and implemented in ANT.Gaussian.[160]

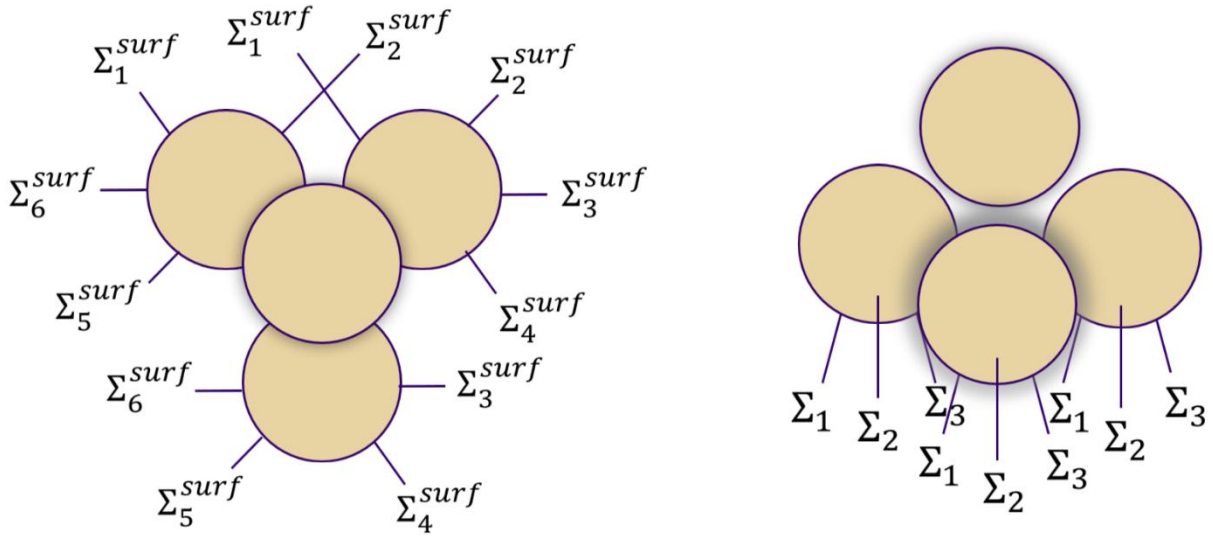


Figure 7.4 - A four atom pyramid contact, showing the top-down view (left) with the surface self-energy contributions and the side view (right) with the bulk self-energy contributions. Each atom corresponds to a row/column in the matrix given by equation 87.

To use the Bethe lattice with a given system, the Fermi energy of the contact is calculated by assuming charge neutrality across each atom in the lattice. This can be done by constructing the Green's function of the bulk contact, which can be generated from equation 85, applying each self-energy to a central bulk site to get

$$G^R(E) = \left(E - i\delta - H - \sum_k \Sigma_j(E) \right)^{-1} \tag{89}$$

where δ is a small positive number used to create the retarded Green's function. This can be integrated as previously discussed using equation 69 and imposing charge neutrality (equation 70) to calculate the Fermi energy of the contact. While the Bethe lattice contact is an energy-dependent self-energy, solving for the bulk and surface self-energies at each energy point using equations 85 and 86, using the contact self-energy matrix calculated at the Bethe lattice Fermi energy can be used to represent the wide-band approximation, as discussed in previous works.[55] That is, if the self-energy matrix does not change significantly across a range of energies near the Fermi energy, the wide-band approximation is reasonable at those energies.

In this work an Au FCC contact was modeled using the Slater-Koster tight binding parameters generated by Papaconstantopolous et al for the 5d, 6s, and 6p orbitals.[163] All contact atoms used the minimal CRENBS basis set,[164] connecting the basis orbitals to the deorthogonalized contact.

Spin Model

All previous equations in this section have been written in the restricted closed-shell form, assuming spin pairing as the default case for device modeling. For the unrestricted open-shell cases, $\hat{F}_{\alpha\beta} = \hat{F}_{\beta\alpha} = 0$, indicating that spin treatment is fully collinear and the spin α direction has arbitrary spatial direction. To add non-collinear effects, the unconstrained generalized open-shell

model was used, which allowed modeling of spatial spin direction as well spin interactions. This allows us to replace spin directions α and β with set spin up and spin down directions, usually aligned to the z-direction (the default in G16). While ferromagnetic contacts were not explored in this work (Fermi energy was set independent of electron spin), spin-orbit coupling was applied to the contact atoms in the device (Figure 7.2) for the Au case using spin-orbit coupling coefficients applied to the pseudopotential. The previously described formulism for spin-dependent transmission was used, given by equation 36.

In our calculation of non-collinear effects, only spin-orbit coupling was included in the calculation using a pseudopotential applied to each atom center to create a spin-orbit splitting, omitting any two electron integrals required for relativistic calculations.[165] This approximation is applied under the assumption that inter-atomic spin-orbit terms and spin-spin interactions are much smaller. The spin-orbit potential applied to each atomic center is given by a one-electron term given by:

$$\hat{V}_{SO} = \hat{\zeta}_l(\vec{r})\, \hat{\vec{L}} \cdot \hat{\vec{S}} \tag{90}$$

where $\hat{\zeta}_l(\vec{r})$ is a pseudopotential defined for each subshell on the atom center.[166] This is a convenient approximation for larger atoms with non-negligible spin-orbit effects, where core electrons are represented by frozen pseudopotentials in an effective core potential (ECP), and the $\hat{\zeta}_l(\vec{r})$ pseudopotential can be defined as a scaling factor applied to the ECP as a fitting factor based on observed or modeled orbital energy splitting. Using this implementation, any inter-atomic spin-orbit effects are caused by angular momentum coupling between neighboring orbitals or spin

exchange across the molecule. Using this method, basis sets with ECPs can be written to include an averaged relativistic term and a relativistic correction term, which has been parameterized for multiple sparse[164] and larger[167] basis sets.

DFT Setup

The J26 version of the Gaussian Development Version was used for all DFT calculations. For all molecules, a mixed Gaussian basis set was used for atomic orbitals with the hybrid functional for all Fock matrix calculations. For all systems, the organic elements (C, N, O, S, H) used the 6-31G** Pople basis set[137] while the Fe center of the heme used the LANL2DZ basis[60] with scalar relativistic effective core potential (no spin-orbit coefficients applied). For gold atoms not attached to a Bethe lattice contact, the Stuttgart basis set[167] was used, while the minimal CRENBS basis[164] was used for atoms attached to the Bethe lattice for compatibility with the defined Slater-Koster coefficients. Spin orbit coefficients were defined in both Au basis sets. For the modeled heme cases with solvation, the conductor-like polarizable continuum model (C-PCM) was used to model a water dielectric.[168] For all cases, contacts were applied in the z-cartesian direction, to ensure that the S_z spin direction is normal to the contacts. That is, the semi-infinite Bethe lattice lead extended in the -z direction while the triangular tip pointed in the +z direction.

Two functionals were compared for benchmarking of the results, B3LYP[47] and HSE06,[169] which have been used extensively for molecular and solid state systems respectively. The HSE06 functional was shown to reduce the effect of oxidation on the electronic structure while matching

up with the B3LYP case almost exactly in the reduced heme case (see Figure 7.5). For this reason, the B3LYP functional was used for all calculations, to ensure that redox effects were properly accounted for.

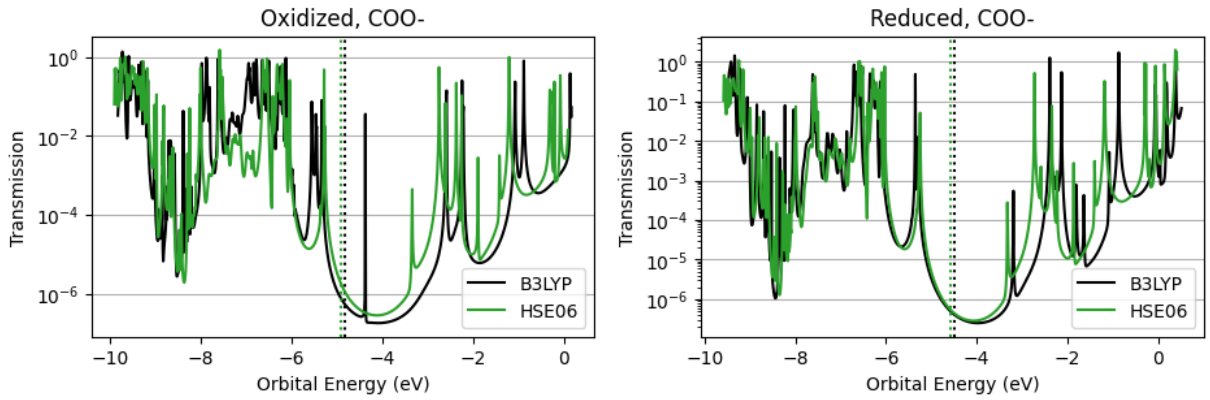


Figure 7.5 - Benchmarking of B3LYP functional against HSE06 hybrid functional for the full Bethe lattice case. In the HSE06 case the two transmission curves don't shift significantly, showing the effect of a lower exchange term.

b. Results

The results from calculations in the PDT system showed that the Bethe lattice contact model matched other modeling and experimental results. Applying the energy-independent self-energy approximation was also demonstrated to be a valid approximation both for the PDT and coordinated heme systems. This approximation significantly sped up calculation time and was used for the expanded heme system to generate results.

Model Verification

To test the effects of contact geometry, four geometries were tested with the system to look at the effect of contact size, unpaired electrons, and overall shape on the transmission. The effect of unpaired electrons was clearly shown in the 4- and 5- atom triangular contacts (see Figure 7.6 below), with the 5-atom contacts shifting the transmission curve significantly. Additionally, constraining the system with a restricted closed-shell model (no net spin, singlet state) and using the charge neutrality boundary condition for Fermi energy calculation, the unpaired orbital in each of the Au contacts induced a spin filter effect in the transmission at the Fermi energy. This occurred despite applying the spin-independent Bethe lattice contact. In each of the other contact geometries, shown in Figure 7.6A, there were an even number of Au atoms (and by extension an even number of electrons), and no spin dependence in the current was observed near the Fermi energy after applying the open-shell model, as shown in Figure 7.6B.

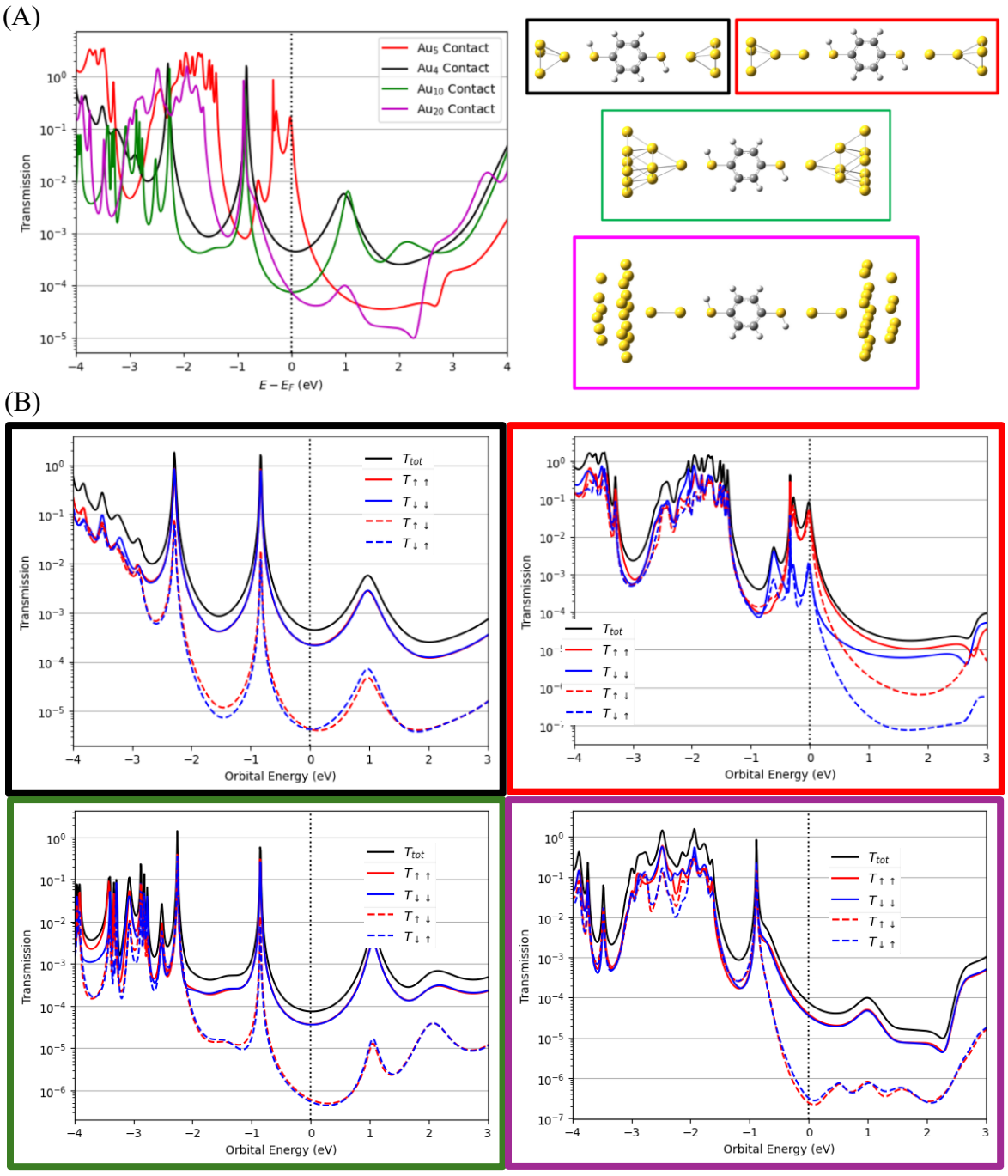


Figure 7.6 – (A) Comparison of total transmission for the Au contact geometries, shown color coded to the right of the plot, followed by (B) color coded spin-dependent transmission plots

Each of the contacts with even numbers of atoms showed similar trends in the transmission, with the total transmission decreasing for the larger extended contact geometries. This trend matches the expected behavior of the changing geometry: increasing the number of contact atoms increases the height of the contact, causing the distance between the simulated Bethe lattice planes of the left and right contact to also increase. Temperature dependence was tested for the Au_4 contact case, though the resulting transmission curve was identical, with the Fermi energy only shifting up by about 0.2eV at room temperature, as shown in Figure 7.7 below.

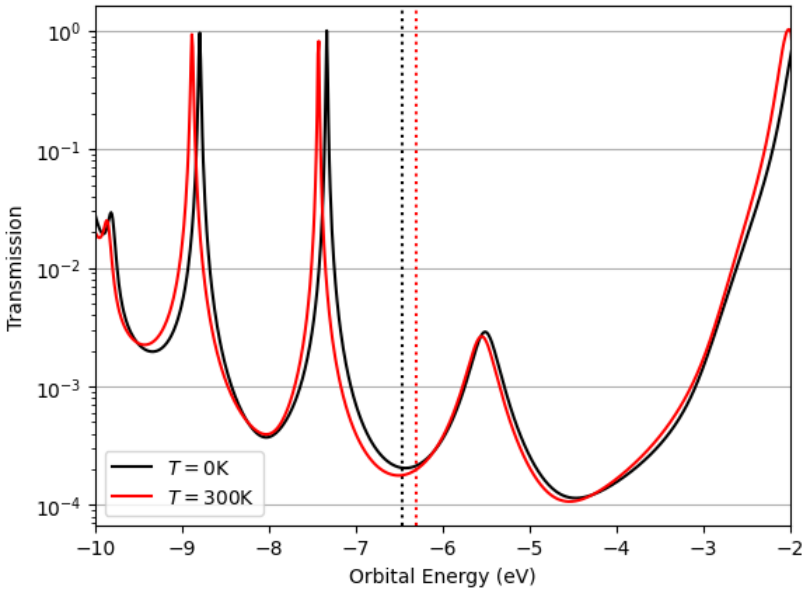


Figure 7.7 - A comparison of the Au4 atom contact for two temperature cases: 300K and 0K. The transmission remains nearly identical for both cases, with the Fermi energy only shifting up by about 0.2eV for the 300K case.

In previous work, the wide-band approximation has been applied to molecular contacts, using a constant self-energy value across all electron energies, which has previously been shown to be a valid approximation for gold contacts.[55] This assumption was tested first on the PDT system and then on the coordinated heme system comparing the energy-independent approximation to the full

energy-dependent Bethe lattice. Transmission curves were compared for each of these cases, shown in the Figure 7.8, and IV characteristics were generated for the PDT system with and without spin-orbit coupling applied, shown in Figure 7.9a. The IV characteristics were generated by applying a bias sweep across the extended system comparing open-shell and closed-shell models, with a further comparison of the energy-dependent Bethe contact and energy-independent contact in the open-shell case. The IV characteristics, as shown in Figure 7.9a, line up at all applied voltages, with only small deviations shown in the open-shell energy-dependent Bethe lattice model. This appears as a hysteresis in the voltage sweep and can be explained by convergence errors under applied bias, as a lower convergence criterion was required for these voltage points. This further demonstrates the efficiency of the energy-independent approximation, which can stabilize the NEGF-DFT calculation under applied biases, improving the convergence of the calculation and resulting accuracy.

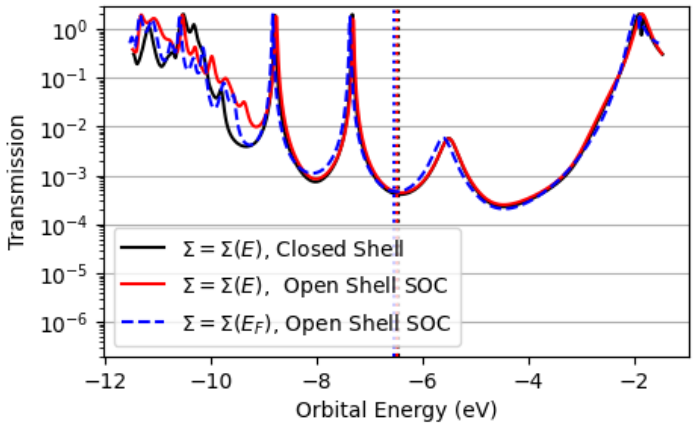


Figure 7.8 - A comparison of the Au4 atom system for three cases: black - energy-dependent contact self-energies and closed-shell model, red - energy-dependent contact self-energies with open-shell model (spin-orbit coupling included), and blue dashed - energy-independent contact self-energies with open-shell model (spin-orbit coupling included)

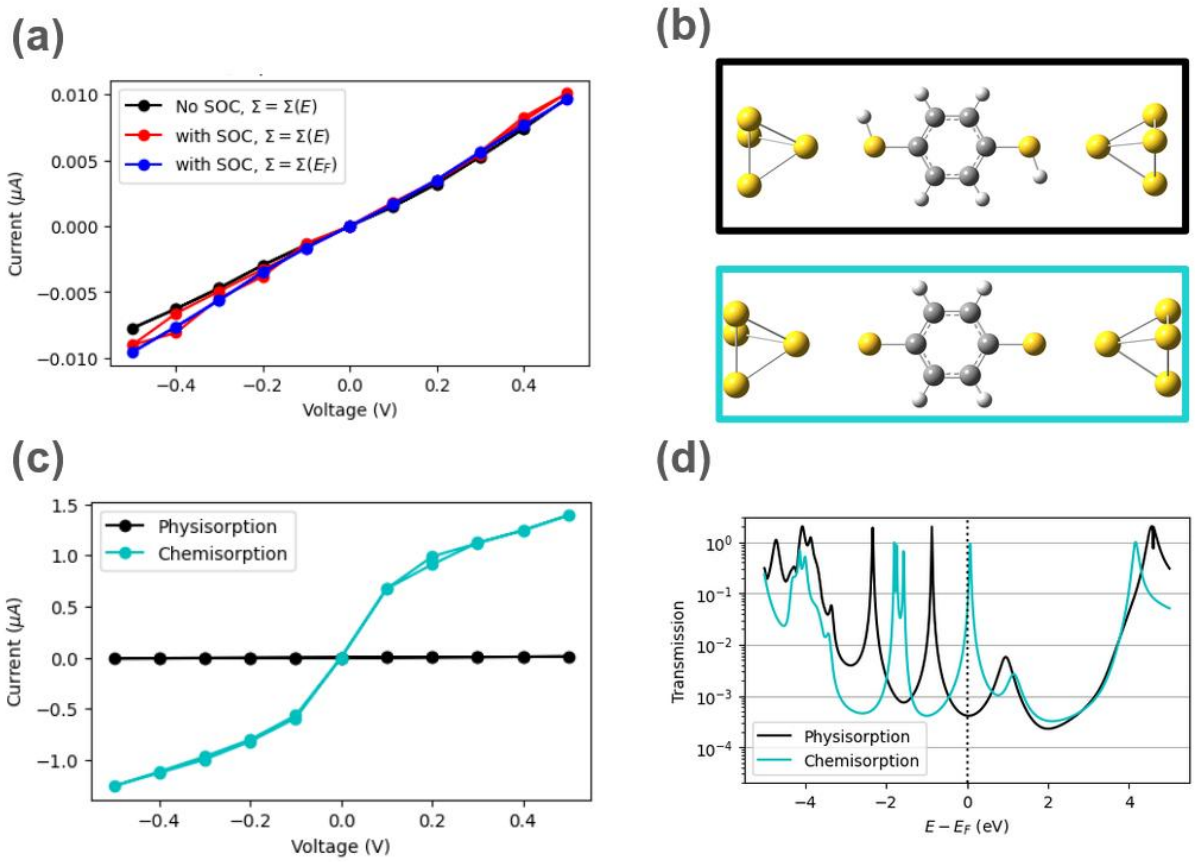


Figure 7.9 - Model verification results with the Phenyl Dithiol (PDT) system: IV characteristics were first collected for the physisorption case with and without spin-orbit coupling as well as with and without the energy-independent self-energy approximation, shown superimposed in (a). This structure was modified to generate the chemisorption case (hydrogens removed) as shown in (b), with the IV characteristics comparing the two systems are shown in (c) and comparison of transmission at equilibrium shown in (d).

PDT Experimental Validation

To evaluate the validity of our computational approach, we compared our modeled current-voltage (IV) characteristics, shown in Figure 7.9c, to both previously modeled results and experimentally measured conductance values. The full Bethe lattice was applied to the 4-Au atom contact using these contact geometries to compare scenarios with and without thiol bonding, with

structures shown in Figure 7.9b. We distinguish between two cases: the chemisorption case, where the sulfur-bonded hydrogens are removed, and the physisorption case, where the hydrogens remain bonded to the sulfur. In both cases, all other atomic coordinates remain the same. As shown in Figure 7.9d, transmission increases significantly for the chemisorption case, matching up with previous modeling results.[148,149] Though the atomic coordinates are the same for both cases except for the removal of the hydrogens (the gold contacts remain 2.8 Å away from the sulfur groups), the conductivity was determined to be 0.06 conductance quanta (G_0), which falls right within the range of values measured experimentally (0.001-0.1 G_0).[150,152] As a well characterized and studied molecular junction, this provides a experimentally grounded benchmark for further modeling results.

Coordinated Heme Results

To benchmark the energy-independent contact approximation against the full energy-dependent Bethe lattice, the smaller 4-atom contact system was compared both for the oxidized Fe^{III} and reduced Fe^{II} cases. The transmission curves line up well across energies, with the Fermi energies shifting between the cases by up to 0.5eV, shown in Figure 7.10 below. Therefore, for a more accurate analysis, spin effects were analyzed on the 10-atom contact. Not only did the 10-atom contact have a flatter transmission curve, making it less dependent on the specified Fermi energy, but the added basis functions on the second layer of gold atoms allowed for less constraints on the contact-molecule interface.

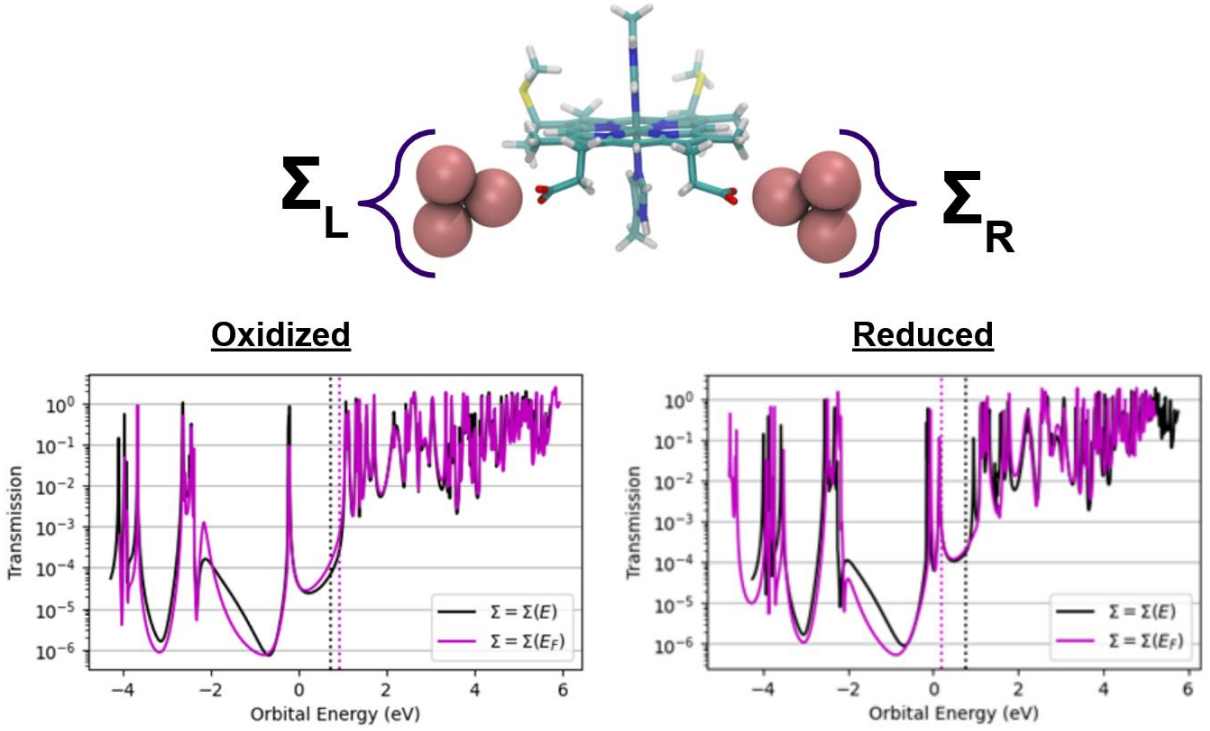


Figure 7.10 - Comparison of the generated coherent transmission plots generated using energy-independent and energy-dependent self-energy applied to the 4-Au contact.

The full 10-atom contact showed no spin dependence in transmission at or near the Fermi energy, as shown in Figure 7.12. This result extends to the protonated propionate case, which would appear in an oxidizing environment, showing that this result isn't sensitive to the excess charge or modified chemistry of the interface as seen with the PDT molecule. The protonated case was also used to check the effect of the solvent model, which was shown only to shift the Fermi energy as shown in Figure 7.11 (only the oxidized case showed convergence without solvent). Analysis of the spin density, plotted below the transmission graphs in Figure 7.12, showed spin accumulation on the Fe in the reduced case, both with and without propionate protonation, and in

the oxidized case with protonation (COO-). Note that the convergence criteria for each of these cases was 1E-3 except for the oxidized case without protonation (COO-), which had a final convergence value of 1.6E-3. Accumulation of spins on the Fe center was also seen in the 4-atom contact reduced cases (both with and without energy dependence) and was not seen in the initial DFT results before NEGF-DFT, for all cases.

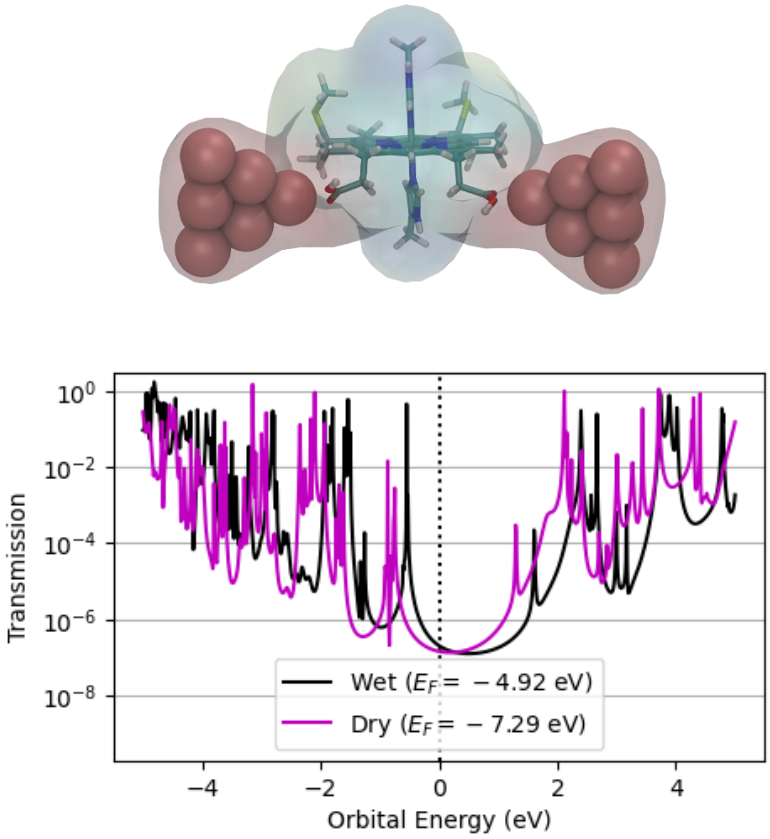


Figure 7.11 - A comparison of wet and dry transmission curves for the oxidized Fe^{III} heme system with hydrogenated COOH groups. The structure is shown above with the solvent surface mesh highlighted.

.

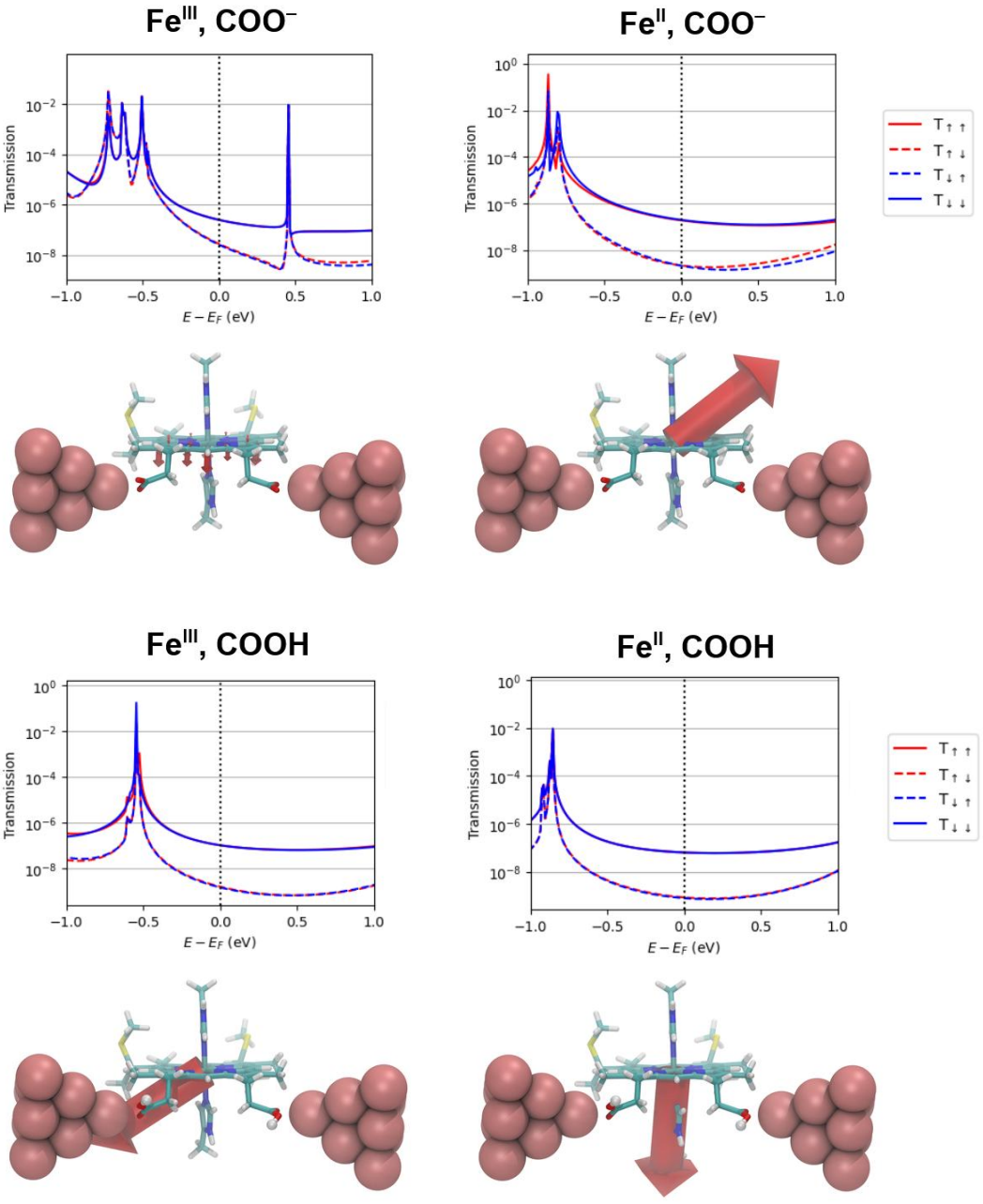


Figure 7.12 - Spin-dependent transmission and spin density maps for the four cases: oxidized (left column), reduced (right column), dehydrogenated propionate group (top row) and hydrogenated propionate group (bottom row). Transmission plots are plotted for energies within 1eV of the Fermi energy (E_F) and spin density plots assign spin

vectors to each atom based on the sum of the expectation value of spin in each direction for each atomic basis function. The vectors (red arrows) are scaled to a length given by $\frac{20\,\langle S_i \rangle}{\hbar}$ Å *for each spin direction.*

c. Discussion

Using a newly developed setup to model non-collinear spin, we demonstrated that the contact induces spin accumulation on the heme molecule. Many of the underlying assumptions in this model were checked with generated results or comparing to previous work, and an exhaustive list is given in the following table of assumptions:

Assumption	Physical meaning	Impact on Results
Ballistic limit for quantum transport	All transport calculations were done in the ballistic limit, not including decoherence or interactions with environment	System sizes were significantly lower than the decoherence lengths expected, with high coupling between all atoms. We suspect these effects are negligible
Spin-independent contacts	Bethe lattice self-energy is block diagonal, repeating the same values for spin up and spin down orbitals	All spin-orbit effects were limited to extended atoms in the system. This allowed us to focus specifically on this interface, but bulk properties may induce spin-dependent effects on the contact.
η = 1e-10 eV	This minimal broadening parameter was used to converge retarded surface Green's function	The imaginary part of the self-energy was significantly higher than eta in all cases, we suspect this has no impact on the results.
Implicit Solvation around heme molecule	The conductive polarizable continuum model was used to model the dielectric behavior of the water and protein backbone.	This assumption has been proven to be valid in previous studies.[170] Removing solvent shifted all orbital energies equally (see Figure 7.11)
Minimal basis set for contact atoms	All Au atoms attached to Bethe lattices use a minimal basis (one basis per atomic orbital) as specified by the CRENBS basis[171]	A comparison of the full Stuttgart basis[167] vs CRENBS basis was shown not to change the result for the PDT system with the PDT atom

Tight-binding model for Au contacts	To limit computational intensity, previously derived tight-binding coefficients were used for modeling the semi-infinite contacts	This assumption was based on previous work that showed that the Bethe lattice was a valid approximation for metal contacts in molecular devices[160]
System Fermi energy calculated from available electrons	The Fermi energy of the extended system was calculated based on the number of electrons from the Gaussian input file. The Fermi energy of the contacts, calculated based on electroneutrality, were adjusted to match this energy for all NEGF calculations	While this approximation limits the study of partial charges induced by the contacts on the device, which may be possible with robust Fermi energy modeling of the contact lattice, including a large contact in the extended system limits this effect on the system.

While the transmission itself shows no filtering effects, modeling the Au surface in contact with the heme molecule leads to spin accumulation on the Fe center. This occurs without the addition of an external magnetic field and is purely caused by spin-orbit interactions with the gold orbitals. This result was confirmed by changing the contact chemistry (hydrogenation of the propionate group) and the solvation model applied and was benchmarked between two functionals. Of these cases, the reduced case without hydrogenation (COO-) and the oxidized case with hydrogenation (COOH) were the most stable cases for the NEGF-DFT solver, reaching convergence quickly after less than 50 cycles. In all cases, the energy-independent self-energy approximation reduced computation time by several orders of magnitude, allowing for only one matrix inversion operation per cycle rather than the hundreds required for direct integration at the expense of some stability in convergence.

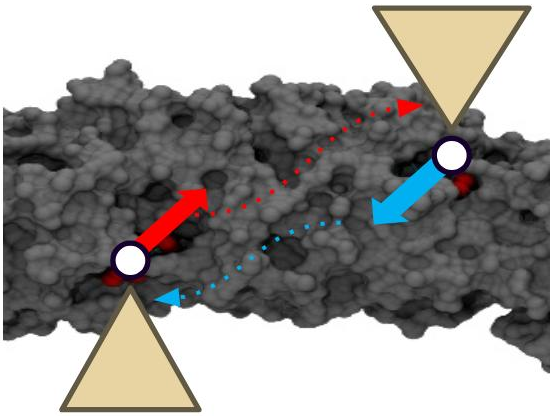


Figure 7.13 - A schematic of a mechanism of spin filtering induced by the contacts. Red arrows show spin filtering of electrons injected from the bottom contact and blue arrows show spin filtering of electrons injected from the top contact.

While spin filtering through the heme was not observed near the Fermi energy in any of the heme setups, the spin accumulation can begin to explain experimental results that show spin selectivity. Further spin filtering, either by quantum tunneling or hopping mechanisms, would be induced by initial contact geometry effects, as shown in Figure 7.13. This result is purely caused by the spin-orbit coupling of the nearest Au cluster in the extended system and does not include the Rashba spin-orbit effect caused by the full Au surface, which can create surface spin states at the cytochrome-contact interface.[140] This effect is key to explaining the large magnetoresistance seen by cytochromes deposited on an Au surface, especially those seen across length scales smaller than the helicity of the cytochrome structure. The simulation results from this work extend to experimental setups with ferromagnetic filtering,[127] which typically use a gold layer for contact with the cytochrome system to prevent oxidation of the ferromagnetic surface. In such setups, our results predict that the contact interface will contribute to the spin filtering effect measured.

8. Future Work

In these works, multiple new models have been developed for modeling of modified DNA molecules and cytochrome proteins, validating each new finding with previous theory or experimental results. The modeling results have demonstrated the importance of the transition metal centers in each system, which are central for charge transport in each of the studied systems. Before expanding these models, it would first be useful to apply them to additional systems and verify with experiment. In each case, additional modeling runs with similar systems would help understand any limits of the models and check any underlying assumptions. After benchmarking, expanding the GauNEGF package to run more efficiently and include additional effects would be helpful for future modeling and materials development, especially as it is applied to mmDNA systems.

a. Benchmarking With Experiment

For the metal modified DNA system, the next steps would involve modeling larger structures and looking at the effect on electronic structure. Applying the textbook DNA structure for an mmDNA wire (each base pair is metallated) results in high transmission with multiple bands, but a more thorough modeling of the structure may change these results. Currently this model is being applied to the commonly used C-Ag-C and T-Hg-T basepair systems, and future work will expand it to the other basepairs in the library. Another consideration for these systems, especially with

heavier atoms like Hg, is the effect of spin-orbit coupling, which may induce spin procession during transport. This could be modeled using open-shell methods and compared to experimental setups that measure spin filtering. Additionally, the impact of physical electrical contacts using the NEGF-DFT model will be helpful for comparing to break-junction experiments. Ideally, using the force fields and NEGF-DFT modeling system, it could be possible to fully model the dynamics and currents observed during break junction or scanning probe microscopy experiments.

For the cytochrome system, it would be helpful to apply the NEGF-DFT model to other contact interfaces and compare more directly with experiment. One specific example is the Zn-modified heme site, which has been experimentally observed to remove any spin dependence. Additionally, it would be helpful to model a multi-heme system between two contacts to look at the effects of inter-heme geometry on the electron and spin transport. This could be compared to experimental results from several multi-heme cytochromes including STC,[128] MtrA,[127] OmcS,[5] and OmcZ.[147] Working with experimentalists, it would be possible to look at potential applications of the spin transport properties of these cytochromes, such as the spin-dependent oxygen evolution reaction that can be used for green hydrogen generation.[172] Such applications would be directly helpful for a transition to renewable energy.[173]

b. Expanding GauNEGF Package

Building on the GauNEGF package, there are two major directions to go: (a) improve calculation efficiency and (b) expand model to include environmental effects. The first direction is crucial: as it stands, the NEGF-DFT calculation is quite intensive and requires many large matrix operations especially for energy-dependent self-energy. These energy-dependent calculations are especially important for the second direction, because this mode is required for more realistic contact models and models with finite temperature and decoherence effects. A first step for increasing efficiency is benchmarking the parallelization available using the optimization built into NumPy and similar packages for graphics cards, like CuPy.[174] This could be done with DNA systems comparing systems of different sizes, and any limitations could be addressed by parallelizing other parts of the code. It may also be helpful to use more of the built-in Gaussian methods, which have been optimized for ab-initio calculations. Using machine learning or other regression methods could be helpful within the calculation, to reach convergence more quickly, or to predict behavior of new systems. Prediction capabilities would be a long term goal after building a database of benchmarked results, much like what is currently done with large DFT datasets like the JARVIS database[175] developed at NIST.

There are several improvements to the GauNEGF code that could be implemented without increasing computational requirements. One major feature missing is modeling of magnetic fields, which can be modeled with the Zeeman splitting, given by

$$\hat{V}_M = \frac{g\mu_B}{\hbar} \hat{\vec{J}} \cdot \hat{\vec{B}}$$

where $\hat{\vec{J}} \cdot \hat{\vec{B}}$ is the dot product of the time averaged total angular momentum J and the magnetic field B, g is the Landé g-factor, and μ_B is the Bohr magneton. In the weak field limit, this can be added as a perturbation to the NEGF-DFT cycle, eventually converging to a self-consistent result. In addition to magnetic fields, the software package can be extended to the modeling of spin-dependent contacts. This can range from ferromagnetic contacts to contacts with spin mixing effects, like spin-orbit coupling. For a simple collinear case, the self-energies must be defined for both spin directions, while the non-collinear case must be a larger matrix with four quadrants, spin mixing effects included in the off-diagonal terms. As a final note, the decoherence models that were previously applied to both the cytochrome and DNA systems have not yet been implemented in the GauNEGF package. Each of these features could be added to the code without changing the number of matrix operations, but will require careful consideration of the underlying physics to ensure accuracy.

c. A New Generation of Nanodevices

While traditional silicon-based electronics are being pushed to their limits as devices continue to be scaled down,[176] many researchers are turning towards biomolecules, such as DNA origami, to create the next generation of nanodevices.[177,178] The unique properties of mmDNA and cytochromes, as explored in this work, can add the much-needed conductive properties to these

biomolecule-based electronics. Such devices could be used as spintronic devices, using the spin degree of freedom unveiled from the modeling approaches that have been applied to cytochromes. Already the customizability of mmDNA devices has been explored, showing that a DNA-based electrical switch is possible just by changing the local electrical environment.[179] The integration of biomolecules into electronics into existing technologies poses many new issues, especially at the interfaces, and the tools developed in this work, such as the GauNEGF package, can be applied to these new technologies to address these issues in design and scale-up. By unraveling the complex quantum behaviors that drive the electrical behavior of cytochromes and mmDNA, this work lays essential groundwork for a new generation of bioelectronic technologies that may ultimately transcend the limitations of traditional semiconductor architectures.

References:

(1) Lovley, D. R. E-Biologics: Fabrication of Sustainable Electronics with "Green" Biological Materials. *mBio* **2017**, *8* (3). https://doi.org/10.1128/mBio.00695-17.

(2) Battersby, A. R.; McDonald, E. Origin of the Pigments of Life: The Type-III Problem in Porphyrin Biosynthesis. *Acc. Chem. Res.* **1979**, *12* (1), 14–22. https://doi.org/10.1021/ar50133a003.

(3) Zhang, B.; Song, W.; Pang, P.; Lai, H.; Chen, Q.; Zhang, P.; Lindsay, S. Role of Contacts in Long-Range Protein Conductance. *Proc. Natl. Acad. Sci.* **2019**, *116* (13), 5886–5891. https://doi.org/10.1073/pnas.1819674116.

(4) Malvankar, N. S.; Vargas, M.; Nevin, K. P.; Franks, A. E.; Leang, C.; Kim, B. C.; Inoue, K.; Mester, T.; Covalla, S. F.; Johnson, J. P.; Rotello, V. M.; Tuominen, M. T.; Lovley, D. R. Tunable Metallic-like Conductivity in Microbial Nanowire Networks. *Nat. Nanotechnol.* **2011**, *6* (9), 573–579. https://doi.org/10.1038/nnano.2011.119.

(5) Dahl, P. J.; Yi, S. M.; Gu, Y.; Acharya, A.; Shipps, C.; Neu, J.; Patrick O'Brien, J.; Morzan, U. N.; Chaudhuri, S.; Guberman-Pfeffer, M. J.; Vu, D.; Yalcin, S. E.; Batista, V. S.; Malvankar, N. S. A 300-Fold Conductivity Increase in Microbial Cytochrome Nanowires Due to Temperature-Induced Restructuring of Hydrogen Bonding Networks. *Sci. Adv.* **2022**, *8* (19). https://doi.org/10.1126/sciadv.abm7193.

(6) Nangreave, J.; Han, D.; Liu, Y.; Yan, H. DNA Origami: A History and Current Perspective. *Curr. Opin. Chem. Biol.* **2010**, *14* (5), 608–615. https://doi.org/10.1016/j.cbpa.2010.06.182.

(7) Bouamaied, I.; Nguyen, T.; Rühl, T.; Stulz, E. Supramolecular Helical Porphyrin Arrays Using DNA as a Scaffold. *Org. Biomol. Chem.* **2008**, *6* (21), 3888. https://doi.org/10.1039/b813584c.

(8) Vecchioni, S.; Capece, M. C.; Toomey, E.; Nguyen, L.; Ray, A.; Greenberg, A.; Fujishima, K.; Urbina, J.; Paulino-Lima, I. G.; Pinheiro, V.; Shih, J.; Wessel, G.; Wind, S. J.; Rothschild, L. Construction and Characterization of Metal Ion-Containing DNA Nanowires for Synthetic Biology and Nanotechnology. *Sci. Rep.* **2019**, *9* (1), 6942. https://doi.org/10.1038/s41598-019-43316-1.

(9) Hoshika, S.; Leal, N. A.; Kim, M.-J.; Kim, M.-S.; Karalkar, N. B.; Kim, H.-J.; Bates, A. M.; Watkins, N. E.; SantaLucia, H. A.; Meyer, A. J.; DasGupta, S.; Piccirilli, J. A.; Ellington, A. D.; SantaLucia, J.; Georgiadis, M. M.; Benner, S. A. Hachimoji DNA and RNA: A Genetic System with Eight Building Blocks. *Science* **2019**, *363* (6429), 884–887. https://doi.org/10.1126/science.aat0971.

(10) Fendt, L.-A.; Bouamaied, I.; Thöni, S.; Amiot, N.; Stulz, E. DNA as Supramolecular Scaffold for Porphyrin Arrays on the Nanometer Scale. *J. Am. Chem. Soc.* **2007**, *129* (49), 15319–15329. https://doi.org/10.1021/ja075711c.

(11) Marcus, R. A. Electron Transfer Past and Future; 2007; pp 1–6. https://doi.org/10.1002/9780470141656.ch1.
(12) Kaila, V. R. I.; Johansson, M. P.; Sundholm, D.; Wikström, M. Interheme Electron Tunneling in Cytochrome c Oxidase. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (50), 21470–21475. https://doi.org/10.1073/pnas.1005889107.
(13) Marcus, R. A.; Sutin, N. Electron Transfers in Chemistry and Biology. *Biochim. Biophys. Acta BBA - Rev. Bioenerg.* **1985**, *811* (3), 265–322. https://doi.org/10.1016/0304-4173(85)90014-X.
(14) Jiang, X.; Futera, Z.; Ali, M. E.; Gajdos, F.; Von Rudorff, G. F.; Carof, A.; Breuer, M.; Blumberger, J. Cysteine Linkages Accelerate Electron Flow through Tetra-Heme Protein STC. *J. Am. Chem. Soc.* **2017**, *139* (48), 17237–17240. https://doi.org/10.1021/jacs.7b08831.
(15) Datta, S. *Electronic Transport in Mesoscopic Systems*; Cambridge University Press, 1995. https://doi.org/10.1017/CBO9780511805776.
(16) Kubař, T.; Gutiérrez, R.; Kleinekathöfer, U.; Cuniberti, G.; Elstner, M. Modeling Charge Transport in DNA Using Multi-Scale Methods. *Phys. Status Solidi B* **2013**, *250* (11), 2277–2287. https://doi.org/10.1002/pssb.201349148.
(17) Adessi, Ch.; Walch, S.; Anantram, M. P. Environment and Structure Influence on DNA Conduction. *Phys. Rev. B* **2003**, *67* (8), 081405. https://doi.org/10.1103/PhysRevB.67.081405.
(18) Lebedev, N.; Griva, I.; Blom, A.; Tender, L. M. Effect of Iron Doping on Protein Molecular Conductance. *Phys. Chem. Chem. Phys.* **2018**, *20* (20), 14072–14081. https://doi.org/10.1039/C8CP00656C.
(19) Matsuura, Y.; Kato, F.; Okita, M.; Tachikawa, T. Coherent Spin Transport in a Natural Helical Protein Molecule. *Chem. Phys.* **2021**, *544*, 111107. https://doi.org/10.1016/j.chemphys.2021.111107.
(20) Ru, X.; Zhang, P.; Beratan, D. N. Assessing Possible Mechanisms of Micrometer-Scale Electron Transfer in Heme-Free Geobacter Sulfurreducens Pili. *J. Phys. Chem. B* **2019**, *123* (24), 5035–5047. https://doi.org/10.1021/acs.jpcb.9b01086.
(21) Eshel, Y.; Peskin, U.; Amdursky, N. Coherence-Assisted Electron Diffusion across the Multi-Heme Protein-Based Bacterial Nanowire. *Nanotechnology* **2020**, *31* (31), 314002. https://doi.org/10.1088/1361-6528/ab8767.
(22) Wilner, E. Y.; Wang, H.; Thoss, M.; Rabani, E. Nonequilibrium Quantum Systems with Electron-Phonon Interactions: Transient Dynamics and Approach to Steady State. *Phys. Rev. B* **2014**, *89* (20), 205129. https://doi.org/10.1103/PhysRevB.89.205129.
(23) Wang, J.-S.; Ni, X.; Jiang, J.-W. Molecular Dynamics with Quantum Heat Baths: Application to Nanoribbons and Nanotubes. *Phys. Rev. B* **2009**, *80* (22), 224302. https://doi.org/10.1103/PhysRevB.80.224302.
(24) Gutiérrez, R.; Mandal, S.; Cuniberti, G. Quantum Transport through a DNA Wire in a Dissipative Environment. *Nano Lett.* **2005**, *5* (6), 1093–1097. https://doi.org/10.1021/nl050623g.

(25) Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All Spd-Block Elements (Z = 1–86). *J. Chem. Theory Comput.* **2017**, *13* (5), 1989–2009. https://doi.org/10.1021/acs.jctc.7b00118.
(26) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended <scp>tight-binding</Scp> Quantum Chemistry Methods. *WIREs Comput. Mol. Sci.* **2021**, *11* (2). https://doi.org/10.1002/wcms.1493.
(27) Mohammad, H.; Demir, B.; Akin, C.; Luan, B.; Hihath, J.; Oren, E. E.; Anantram, M. P. Role of Intercalation in the Electrical Properties of Nucleic Acids for Use in Molecular Electronics. *Nanoscale Horiz.* **2021**, *6* (8), 651–660. https://doi.org/10.1039/D1NH00211B.
(28) Qi, J.; Edirisinghe, N.; Rabbani, M. G.; Anantram, M. P. Unified Model for Conductance through DNA with the Landauer-Büttiker Formalism. *Phys. Rev. B* **2013**, *87* (8), 085404. https://doi.org/10.1103/PhysRevB.87.085404.
(29) Reed, M. A.; Zhou, C.; Muller, C. J.; Burgin, T. P.; Tour, J. M. Conductance of a Molecular Junction. *Science* **1997**, *278* (5336), 252–254. https://doi.org/10.1126/science.278.5336.252.
(30) Kergueris, C.; Bourgoin, J.-P.; Palacin, S.; Esteve, D.; Urbina, C.; Magoga, M.; Joachim, C. Electron Transport through a Metal-Molecule-Metal Junction. *Phys. Rev. B* **1999**, *59* (19), 12505–12513. https://doi.org/10.1103/PhysRevB.59.12505.
(31) Emberly, E. G.; Kirczenow, G. Molecular Spintronics: Spin-Dependent Electron Transport in Molecular Wires. *Chem. Phys.* **2002**, *281* (2–3), 311–324. https://doi.org/10.1016/S0301-0104(02)00566-9.
(32) Waldron, D.; Haney, P.; Larade, B.; MacDonald, A.; Guo, H. Nonlinear Spin Current and Magnetoresistance of Molecular Tunnel Junctions. *Phys. Rev. Lett.* **2006**, *96* (16), 166804. https://doi.org/10.1103/PhysRevLett.96.166804.
(33) Petta, J. R.; Slater, S. K.; Ralph, D. C. Spin-Dependent Transport in Molecular Tunnel Junctions. *Phys. Rev. Lett.* **2004**, *93* (13), 136601. https://doi.org/10.1103/PhysRevLett.93.136601.
(34) Dednam, W.; Sabater, C.; Tal, O.; Palacios, J. J.; Botha, A. E.; Caturla, M. J. Refined Electron-Spin Transport Model for Single-Element Ferromagnetic Systems: Application to Nickel Nanocontacts. *Phys. Rev. B* **2020**, *102* (24), 245415. https://doi.org/10.1103/PhysRevB.102.245415.
(35) Evers, F.; Aharony, A.; Bar-Gill, N.; Entin-Wohlman, O.; Hedegård, P.; Hod, O.; Jelinek, P.; Kamieniarz, G.; Lemeshko, M.; Michaeli, K.; Mujica, V.; Naaman, R.; Paltiel, Y.; Refaely-Abramson, S.; Tal, O.; Thijssen, J.; Thoss, M.; van Ruitenbeek, J. M.; Venkataraman, L.; Waldeck, D. H.; Yan, B.; Kronik, L. Theory of Chirality Induced Spin Selectivity: Progress and Challenges. *Adv. Mater.* **2022**, *34* (13), 2106629. https://doi.org/10.1002/adma.202106629.
(36) Xie, Z.; Markus, T. Z.; Cohen, S. R.; Vager, Z.; Gutierrez, R.; Naaman, R. Spin Specific Electron Conduction through DNA Oligomers. *Nano Lett.* **2011**, *11* (11), 4652–4655. https://doi.org/10.1021/nl2021637.

(37) Khvostichenko, D.; Choi, A.; Boulatov, R. Density Functional Theory Calculations of the Lowest Energy Quintet and Triplet States of Model Hemes: Role of Functional, Basis Set, and Zero-Point Energy Corrections. *J. Phys. Chem. A* **2008**, *112* (16), 3700–3711. https://doi.org/10.1021/jp076979t.
(38) Mohammad, H. Modeling and Simulation of Charge Transport Through Nanoscale Nucleic Acid Structures, University of Washington, Seattle, WA, USA, 2022.
(39) Schmidt, A.; Teeter, M.; Weckert, E.; Lamzin, V. S. Crystal Structure of Small Protein Crambin at 0.48 Å Resolution. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* **2011**, *67* (4), 424–428. https://doi.org/10.1107/S1744309110052607.
(40) Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159* (1), 98–103. https://doi.org/10.1103/PhysRev.159.98.
(41) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545–1614. https://doi.org/10.1002/jcc.21287.
(42) Breuer, M.; Zarzycki, P.; Blumberger, J.; Rosso, K. M. Thermodynamics of Electron Flow in the Bacterial Deca-Heme Cytochrome MtrF. *J. Am. Chem. Soc.* **2012**, *134* (24), 9868–9871. https://doi.org/10.1021/ja3027696.
(43) Groenhof, G. Introduction to QM/MM Simulations; 2013; pp 43–66. https://doi.org/10.1007/978-1-62703-017-5_3.
(44) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13* (12), 1173–1213. https://doi.org/10.1007/s00894-007-0233-4.
(45) Kato, T. On the Existence of Solutions of the Helium Wave Equation. *Trans. Am. Math. Soc.* **1951**, *70* (2), 212. https://doi.org/10.2307/1990367.
(46) Fock, V. Näherungsmethode Zur Lösung Des Quantenmechanischen Mehrkörperproblems. *Z. Für Phys.* **1930**, *61* (1–2), 126–148. https://doi.org/10.1007/BF01340294.
(47) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652. https://doi.org/10.1063/1.464913.
(48) Diercksen, G.; McWeeny, R. Self-Consistent Perturbation Theory. I. General Formulation and Some Applications. *J. Chem. Phys.* **1966**, *44* (9), 3554–3560. https://doi.org/10.1063/1.1727264.
(49) Koopmans, T. Über Die Zuordnung von Wellenfunktionen Und Eigenwerten Zu Den Einzelnen Elektronen Eines Atoms. *Physica* **1934**, *1* (1–6), 104–113. https://doi.org/10.1016/S0031-8914(34)90011-2.

(50) Roothaan, C. C. J. Self-Consistent Field Theory for Open Shells of Electronic Systems. *Rev. Mod. Phys.* **1960**, *32* (2), 179–185. https://doi.org/10.1103/RevModPhys.32.179.
(51) Hillier, I. H.; Saunders, V. R. A newSCF Procedure and Its Applications Toab Initio Calculations of the States of the Fluorosulphate Radical. *Int. J. Quantum Chem.* **1970**, *4* (5), 503–518. https://doi.org/10.1002/qua.560040505.
(52) McWeeny, R.; Diercksen, G. Self-Consistent Perturbation Theory. II. Extension to Open Shells. *J. Chem. Phys.* **1968**, *49* (11), 4852–4856. https://doi.org/10.1063/1.1669970.
(53) Tsuchimochi, T.; Scuseria, G. E. Communication: ROHF Theory Made Simple. *J. Chem. Phys.* **2010**, *133* (14). https://doi.org/10.1063/1.3503173.
(54) Valatin, J. G. Generalized Hartree-Fock Method. *Phys. Rev.* **1961**, *122* (4), 1012–1020. https://doi.org/10.1103/PhysRev.122.1012.
(55) Verzijl, C. J. O.; Seldenthuis, J. S.; Thijssen, J. M. Applicability of the Wide-Band Limit in DFT-Based Molecular Transport Calculations. *J. Chem. Phys.* **2013**, *138* (9), 094102. https://doi.org/10.1063/1.4793259.
(56) Zöllner, M. S.; Varela, S.; Medina, E.; Mujica, V.; Herrmann, C. Insight into the Origin of Chiral-Induced Spin Selectivity from a Symmetry Analysis of Electronic Transmission. *J. Chem. Theory Comput.* **2020**, *16* (5), 2914–2929. https://doi.org/10.1021/acs.jctc.9b01078.
(57) Toomey, E.; Xu, J.; Vecchioni, S.; Rothschild, L.; Wind, S.; Fernandes, G. E. Comparison of Canonical versus Silver(I)-Mediated Base-Pairing on Single Molecule Conductance in Polycytosine dsDNA. *J. Phys. Chem. C* **2016**, *120* (14), 7804–7809. https://doi.org/10.1021/acs.jpcc.5b11968.
(58) Woloszyn, K.; Vecchioni, S.; Ohayon, Y. P.; Lu, B.; Ma, Y.; Huang, Q.; Zhu, E.; Chernovolenko, D.; Markus, T.; Jonoska, N.; Mao, C.; Seeman, N. C.; Sha, R. Augmented DNA Nanoarchitectures: A Structural Library of 3D Self-Assembling Tensegrity Triangle Variants. *Adv. Mater.* **2022**, *34* (49), 2206876. https://doi.org/10.1002/adma.202206876.
(59) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16 Revision C.01, 2016.
(60) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for K to Au Including the Outermost Core Orbitals. *J. Chem. Phys.* **1985**, *82* (1), 299–310. https://doi.org/10.1063/1.448975.

(61) Petersson, G. A.; Al-Laham, M. A. A Complete Basis Set Model Chemistry. II. Open-shell Systems and the Total Energies of the First-row Atoms. *J. Chem. Phys.* **1991**, *94* (9), 6081–6090. https://doi.org/10.1063/1.460447.
(62) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105* (8), 2999–3094. https://doi.org/10.1021/cr9904009.
(63) Ono, A.; Torigoe, H.; Tanaka, Y.; Okamoto, I. Binding of Metal Ions by Pyrimidine Base Pairs in DNA Duplexes. *Chem. Soc. Rev.* **2011**, *40* (12), 5855. https://doi.org/10.1039/c1cs15149e.
(64) Kondo, J.; Tada, Y.; Dairaku, T.; Hattori, Y.; Saneyoshi, H.; Ono, A.; Tanaka, Y. A Metallo-DNA Nanowire with Uninterrupted One-Dimensional Silver Array. *Nat. Chem.* **2017**, *9* (10), 956–960. https://doi.org/10.1038/nchem.2808.
(65) De, A.; Mohammad, H.; Wang, Y.; Kubendran, R.; Das, A. K.; Anantram, M. P. Modeling and Simulation of DNA Origami Based Electronic Read-Only Memory. In *2022 IEEE 22nd International Conference on Nanotechnology (NANO)*; IEEE, 2022; pp 385–388. https://doi.org/10.1109/NANO54668.2022.9928676.
(66) Seeman, N. C. Nucleic Acid Junctions and Lattices. *J. Theor. Biol.* **1982**, *99* (2), 237–247. https://doi.org/10.1016/0022-5193(82)90002-9.
(67) Stulz, E. Porphyrin-Modified DNA as Construction Material in Supramolecular Chemistry and Nano-Architectonics. *Chimia* **2015**, *69* (11), 678–683. https://doi.org/10.2533/chimia.2015.678.
(68) Aich, P.; Labiuk, S. L.; Tari, L. W.; Delbaere, L. J. T.; Roesler, W. J.; Falk, K. J.; Steer, R. P.; Lee, J. S. *M*-DNA: A Complex between Divalent Metal Ions and DNA Which Behaves as a Molecular Wire1. *J. Mol. Biol.* **1999**, *294* (2), 477–485. https://doi.org/10.1006/jmbi.1999.3234.
(69) Tanaka, K.; Yamada, Y.; Shionoya, M. Formation of Silver(I)-Mediated DNA Duplex and Triplex through an Alternative Base Pair of Pyridine Nucleobases. *J. Am. Chem. Soc.* **2002**, *124* (30), 8802–8803. https://doi.org/10.1021/ja020510o.
(70) Miyake, Y.; Togashi, H.; Tashiro, M.; Yamaguchi, H.; Oda, S.; Kudo, M.; Tanaka, Y.; Kondo, Y.; Sawa, R.; Fujimoto, T.; Machinami, T.; Ono, A. MercuryII-Mediated Formation of Thymine−HgII−Thymine Base Pairs in DNA Duplexes. *J. Am. Chem. Soc.* **2006**, *128* (7), 2172–2173. https://doi.org/10.1021/ja056354d.
(71) Swasey, S. M.; Rosu, F.; Copp, S. M.; Gabelica, V.; Gwinn, E. G. Parallel Guanine Duplex and Cytosine Duplex DNA with Uninterrupted Spines of AgI-Mediated Base Pairs. *J. Phys. Chem. Lett.* **2018**, *9* (22), 6605–6610. https://doi.org/10.1021/acs.jpclett.8b02851.
(72) Vecchioni, S.; Lu, B.; Livernois, W.; Ohayon, Y. P.; Yoder, J. B.; Yang, C.-F.; Woloszyn, K.; Bernfeld, W.; Anantram, M. P.; Canary, J. W.; Hendrickson, W. A.; Rothschild, L. J.; Mao, C.; Wind, S. J.; Seeman, N. C.; Sha, R. Metal-Mediated DNA Nanotechnology in 3D: Structural Library by Templated Diffraction. *Adv. Mater.* **2023**, *35* (29), 2210938. https://doi.org/10.1002/adma.202210938.

(73) Clever, G. H.; Kaul, C.; Carell, T. DNA–Metal Base Pairs. *Angew. Chem. Int. Ed.* **2007**, *46* (33), 6226–6236. https://doi.org/10.1002/anie.200701185.
(74) Kondo, J.; Yamada, T.; Hirose, C.; Okamoto, I.; Tanaka, Y.; Ono, A. Crystal Structure of Metallo DNA Duplex Containing Consecutive Watson–Crick-like T–HgII–T Base Pairs. *Angew. Chem. Int. Ed.* **2014**, *53* (9), 2385–2388. https://doi.org/10.1002/anie.201309066.
(75) Brancolini, G.; Di Felice, R. Electronic Properties of Metal-Modified DNA Base Pairs. *J. Phys. Chem. B* **2008**, *112* (45), 14281–14290. https://doi.org/10.1021/jp806419t.
(76) Voityuk, A. A. Electronic Coupling Mediated by Stacked [Thymine-Hg-Thymine] Base Pairs. *J. Phys. Chem. B* **2006**, *110* (42), 21010–21013. https://doi.org/10.1021/jp062776d.
(77) Šebera, J.; Řeha, D.; Fukal, J.; Sychrovský, V. Interstrand Charge Transport within Metallo-DNA: The Effect Due to Hg(II)- and Ag(I)-Mediated Base Pairs. *J. Phys. Chem. C* **2020**, *124* (13), 7477–7486. https://doi.org/10.1021/acs.jpcc.9b12020.
(78) Wang, Y.; Luan, B.-Q.; Yang, Z.; Zhang, X.; Ritzo, B.; Gates, K.; Gu, L.-Q. Single Molecule Investigation of Ag+ Interactions with Single Cytosine-, Methylcytosine- and Hydroxymethylcytosine-Cytosine Mismatches in a Nanopore. *Sci. Rep.* **2014**, *4* (1), 5883. https://doi.org/10.1038/srep05883.
(79) Chen, X.; Karpenko, A.; Lopez-Acevedo, O. Silver-Mediated Double Helix: Structural Parameters for a Robust DNA Building Block. *ACS Omega* **2017**, *2* (10), 7343–7348. https://doi.org/10.1021/acsomega.7b01089.
(80) Bachmann, J.; Schönrath, I.; Müller, J.; Doltsinis, N. L. Dynamic Structure and Stability of DNA Duplexes Bearing a Dinuclear Hg(II)-Mediated Base Pair. *Molecules* **2020**, *25* (21), 4942. https://doi.org/10.3390/molecules25214942.
(81) Espinosa Leal, L. A.; Karpenko, A.; Swasey, S.; Gwinn, E. G.; Rojas-Cervellera, V.; Rovira, C.; Lopez-Acevedo, O. The Role of Hydrogen Bonds in the Stabilization of Silver-Mediated Cytosine Tetramers. *J. Phys. Chem. Lett.* **2015**, *6* (20), 4061–4066. https://doi.org/10.1021/acs.jpclett.5b01864.
(82) Vecchioni, S.; Ohayon, Y. P.; Hernandez, C.; Hoshika, S.; Mao, C.; Benner, S. A.; Sha, R. Six-Letter DNA Nanotechnology: Incorporation of Z-P Base Pairs into Self-Assembling 3D Crystals. *Nano Lett.* **2024**, *24* (45), 14302–14306. https://doi.org/10.1021/acs.nanolett.4c03949.
(83) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int. J. Quantum Chem.* **1996**, *60* (7), 1271–1277. https://doi.org/10.1002/(SICI)1097-461X(1996)60:7<1271::AID-QUA8>3.0.CO;2-W.
(84) Li, P.; Merz, K. M. Jr. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56* (4), 599–604. https://doi.org/10.1021/acs.jcim.5b00674.
(85) Ohayon, Y. P.; Hernandez, C.; Chandrasekaran, A. R.; Wang, X.; Abdallah, H. O.; Jong, M. A.; Mohsen, M. G.; Sha, R.; Birktoft, J. J.; Lukeman, P. S.; Chaikin, P. M.; Ginell, S. L.; Mao, C.; Seeman, N. C. Designing Higher Resolution Self-Assembled 3D DNA Crystals via Strand Terminus Modifications. *ACS Nano* **2019**, *13* (7), 7957–7965. https://doi.org/10.1021/acsnano.9b02430.

(86) *Transmetalation for DNA-based Molecular Electronics | Nanoscience | ChemRxiv | Cambridge Open Engage*. https://chemrxiv.org/engage/chemrxiv/article-details/676f690881d2151a0227db16 (accessed 2025-01-08).
(87) Livernois, W.; Li, P. Pymsmt + mmDNA, 2025. https://github.com/wliverno/pymsmt.
(88) Peters, M. B.; Yang, Y.; Wang, B.; Füsti-Molnár, L.; Weaver, M. N.; Merz, K. M. Jr. Structural Survey of Zinc-Containing Proteins and Development of the Zinc AMBER Force Field (ZAFF). *J. Chem. Theory Comput.* **2010**, *6* (9), 2935–2947. https://doi.org/10.1021/ct1002626.
(89) Nilsson, K.; Lecerof, D.; Sigfridsson, E.; Ryde, U. An Automatic Method to Generate Force-Field Parameters for Hetero-Compounds. *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *59* (2), 274–289. https://doi.org/10.1107/S0907444902021431.
(90) Besler, B. H.; Merz Jr., K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11* (4), 431–439. https://doi.org/10.1002/jcc.540110404.
(91) Petersson, G. A.; Al-Laham, M. A. A Complete Basis Set Model Chemistry. II. Open-shell Systems and the Total Energies of the First-row Atoms. *J. Chem. Phys.* **1991**, *94* (9), 6081–6090. https://doi.org/10.1063/1.460447.
(92) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for K to Au Including the Outermost Core Orbitals. *J. Chem. Phys.* **1985**, *82* (1), 299–310. https://doi.org/10.1063/1.448975.
(93) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105* (8), 2999–3094. https://doi.org/10.1021/cr9904009.
(94) Sengupta, A.; Li, Z.; Song, L. F.; Li, P.; Merz, K. M. Jr. Parameterization of Monovalent Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *J. Chem. Inf. Model.* **2021**, *61* (2), 869–880. https://doi.org/10.1021/acs.jcim.0c01390.
(95) Li, Z.; Song, L. F.; Li, P.; Kenneth M. Merz, J. Systematic Parametrization of Divalent Metal Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *J. Chem. Theory Comput.* **2020**. https://doi.org/10.1021/acs.jctc.0c00194.
(96) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model*; 1993; Vol. 97, pp 10269–10280. https://pubs.acs.org/sharingguidelines.
(97) We use the toolbox implemented in Amber 23: Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T. J.; Götz, A. W.; Gohlke, H.; Izadi, S.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kurtzman, T.; Lee, T.-S.; Li, P.; Liu, J.; Luchko, T.; Luo, R.; Manathunga, M.; Machado, M. R.; Nguyen, H. M.; O'Hearn, K. A.; Onufriev, A. V.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Risheh, A.; Schott-Verdugo, S.; Shajan, A.; Swails, J.; Wang, J.; Wei, H.; Wu, X.; Wu, Y.; Zhang, S.; Zhao, S.; Zhu, Q.; Cheatham, T. E. I.; Roe, D. R.; Roitberg, A.; Simmerling, C.; York, D. M.; Nagan, M. C.; Merz, K. M. Jr. AmberTools. *J. Chem. Inf. Model.* **2023**, *63* (20), 6183–6191. https://doi.org/10.1021/acs.jcim.3c01153.

(98) van Gunsteren, W. F.; Berendsen, H. J. C. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34* (5), 1311–1327. https://doi.org/10.1080/00268977700102571.
(99) Joung, I. S.; Cheatham, T. E. I. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020–9041. https://doi.org/10.1021/jp8001614.
(100) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.
(101) Ivani, I.; Dans, P. D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; Portella, G.; Battistini, F.; Gelpí, J. L.; González, C.; Vendruscolo, M.; Laughton, C. A.; Harris, S. A.; Case, D. A.; Orozco, M. Parmbsc1: A Refined Force Field for DNA Simulations. *Nat. Methods* **2016**, *13* (1), 55–58. https://doi.org/10.1038/nmeth.3658.
(102) Li, P.; Song, L. F.; Merz, K. M. Jr. Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. *J. Chem. Theory Comput.* **2015**, *11* (4), 1645–1657. https://doi.org/10.1021/ct500918t.
(103) Li, P.; Merz, K. M. Jr. Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J. Chem. Theory Comput.* **2014**, *10* (1), 289–297. https://doi.org/10.1021/ct400751u.
(104) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197. https://doi.org/10.1021/ja00124a002.
(105) Ucisik, M. N.; Chakravorty, D. K.; Merz, K. M. Jr. Structure and Dynamics of the N-Terminal Domain of the Cu(I) Binding Protein CusB. *Biochemistry* **2013**, *52* (39), 6911–6923. https://doi.org/10.1021/bi400606b.
(106) Lu, X.; Olson, W. K. 3DNA: A Software Package for the Analysis, Rebuilding and Visualization of Three-dimensional Nucleic Acid Structures. *Nucleic Acids Res.* **2003**, *31* (17), 5108–5121. https://doi.org/10.1093/nar/gkg680.
(107) Gusarova, E. B.; Kovaleva, N. A. Molecular Dynamics Simulation of Melting of the DNA Duplex with Silver-Mediated Cytosine–Cytosine Base Pair. *Computation* **2024**, *12* (7), 145. https://doi.org/10.3390/computation12070145.
(108) Bond, D. R.; Lovley, D. R. Electricity Production by *Geobacter Sulfurreducens* Attached to Electrodes. *Appl. Environ. Microbiol.* **2003**, *69* (3), 1548–1555. https://doi.org/10.1128/AEM.69.3.1548-1555.2003.
(109) Wang, F.; Gu, Y.; O'Brien, J. P.; Yi, S. M.; Yalcin, S. E.; Srikanth, V.; Shen, C.; Vu, D.; Ing, N. L.; Hochbaum, A. I.; Egelman, E. H.; Malvankar, N. S. Structure of Microbial Nanowires Reveals Stacked Hemes That Transport Electrons over Micrometers. *Cell* **2019**, *177* (2), 361-369.e10. https://doi.org/10.1016/j.cell.2019.03.029.

(110) Leys, D.; Meyer, T. E.; Tsapin, A. S.; Nealson, K. H.; Cusanovich, M. A.; Van Beeumen, J. J. Crystal Structures at Atomic Resolution Reveal the Novel Concept of "Electron-Harvesting" as a Role for the Small Tetraheme Cytochrome *c* *. *J. Biol. Chem.* **2002**, *277* (38), 35703–35711. https://doi.org/10.1074/jbc.M203866200.

(111) Breuer, M.; Rosso, K. M.; Blumberger, J. Electron Flow in Multiheme Bacterial Cytochromes Is a Balancing Act between Heme Electronic Interaction and Redox Potentials. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (2), 611–616. https://doi.org/10.1073/pnas.1316156111.

(112) Engel, G. S.; Calhoun, T. R.; Read, E. L.; Ahn, T. K.; Mančal, T.; Cheng, Y. C.; Blankenship, R. E.; Fleming, G. R. Evidence for Wavelike Energy Transfer through Quantum Coherence in Photosynthetic Systems. *Nature* **2007**, *446* (7137), 782–786. https://doi.org/10.1038/nature05678.

(113) Leitão, E. F. V.; Ventura, E.; de Santana, O. L.; do Monte, S. A. Electronic Properties of the Low-Lying Spin States of Dimethylnitrosamine Coordinated to Fe(III) Heme Models: An Ab Initio Study. *Int. J. Quantum Chem.* **2014**, *114* (8), 508–520. https://doi.org/10.1002/qua.24595.

(114) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1–3), 51–57. https://doi.org/10.1016/j.cplett.2004.06.011.

(115) Livernois, W.; Anantram, M. P. Quantum Transport in Conductive Bacterial Nanowires. In *2021 IEEE 16th Nanotechnology Materials and Devices Conference, NMDC 2021*; IEEE, 2021; pp 1–5. https://doi.org/10.1109/NMDC50713.2021.9677490.

(116) Wilhelm, J.; Walz, M.; Evers, F. Ab Initio Spin-Flip Conductance of Hydrogenated Graphene Nanoribbons: Spin-Orbit Interaction and Scattering with Local Impurity Spins. *Phys. Rev. B* **2015**, *92* (1), 014405. https://doi.org/10.1103/PhysRevB.92.014405.

(117) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197. https://doi.org/10.1021/ja00124a002.

(118) Huheey, J. E.; Keiter, E. A.; Keiter, R. L. *Inorganic Chemistry: Principles of Structure and Reactivity*; Harper Collins: New York, USA, 1993.

(119) Niu, J.; Dong, S. An Integrated Calcium Fluoride Crystal Ir Thin-Layer Cell and Its Application to Identification of Electrochemical Reduction Product of Bilirubin. *Electrochimica Acta* **1995**, *40* (7), 823–828. https://doi.org/10.1016/0013-4686(94)00369-C.

(120) Smith, D. M. A.; Dupuis, M.; Vorpagel, E. R.; Straatsma, T. P. Characterization of Electronic Structure and Properties of a Bis(Histidine) Heme Model Complex. *J. Am. Chem. Soc.* **2003**, *125* (9), 2711–2717. https://doi.org/10.1021/ja0280473.

(121) Foster, J. P.; Weinhold, F. Natural Hybrid Orbitals. *J. Am. Chem. Soc.* **1980**, *102* (24), 7211–7218. https://doi.org/10.1021/ja00544a007.

(122) Garg, K.; Ghosh, M.; Eliash, T.; Wonderen, J. H. van; N. Butt, J.; Shi, L.; Jiang, X.; Zdenek, F.; Blumberger, J.; Pecht, I.; Sheves, M.; Cahen, D. Direct Evidence for Heme-Assisted Solid-State Electronic Conduction in Multi-Heme c -Type Cytochromes. *Chem. Sci.* **2018**, *9* (37), 7304–7310. https://doi.org/10.1039/C8SC01716F.
(123) Grabert, Hermann, M. H. D. Single Charge Tunneling, Coulomb Blockade Phenomena in Nanostructures. *Science* **1993**.
(124) Xiong, Z. H.; Wu, D.; Valy Vardeny, Z.; Shi, J. Giant Magnetoresistance in Organic Spin-Valves. *Nature* **2004**, *427* (6977), 821–824. https://doi.org/10.1038/nature02325.
(125) Futera, Z.; Wu, X.; Blumberger, J. Tunneling-to-Hopping Transition in Multiheme Cytochrome Bioelectronic Junctions. *J. Phys. Chem. Lett.* **2023**, 445–452. https://doi.org/10.1021/acs.jpclett.2c03361.
(126) Guberman-Pfeffer, M. J. To Be or Not to Be a Cytochrome: Electrical Characterizations Are Inconsistent with Geobacter Cytochrome 'Nanowires.' *Front. Microbiol.* **2024**, *15*. https://doi.org/10.3389/fmicb.2024.1397124.
(127) Mishra, S.; Pirbadian, S.; Mondal, A. K.; El-Naggar, M. Y.; Naaman, R. Spin-Dependent Electron Transport through Bacterial Cell Surface Multiheme Electron Conduits. *J. Am. Chem. Soc.* **2019**, *141* (49), 19198–19202. https://doi.org/10.1021/jacs.9b09262.
(128) Niman, C. M.; Sukenik, N.; Dang, T.; Nwachukwu, J.; Thirumurthy, M. A.; Jones, A. K.; Naaman, R.; Santra, K.; Das, T. K.; Paltiel, Y.; Baczewski, L. T.; El-Naggar, M. Y. Bacterial Extracellular Electron Transfer Components Are Spin Selective. *J. Chem. Phys.* **2023**, *159* (14), 145101. https://doi.org/10.1063/5.0154211.
(129) Livernois, W.; Anantram, M. P. A Spin-Dependent Model for Multi-Heme Bacterial Nanowires. *ACS Nano* **2023**, *17* (10), 9059–9068. https://doi.org/10.1021/acsnano.2c12027.
(130) Matsuura, Y. Coherent Spin Transport in a Multi-Heme Protein Molecule. *Chem. Phys.* **2022**, *558* (March), 111510. https://doi.org/10.1016/j.chemphys.2022.111510.
(131) Maeda, K.; Robinson, A. J.; Henbest, K. B.; Hogben, H. J.; Biskup, T.; Ahmad, M.; Schleicher, E.; Weber, S.; Timmel, C. R.; Hore, P. J. Magnetically Sensitive Light-Induced Reactions in Cryptochrome Are Consistent with Its Proposed Role as a Magnetoreceptor. *Proc. Natl. Acad. Sci.* **2012**, *109* (13), 4774–4779. https://doi.org/10.1073/pnas.1118959109.
(132) Mouesca, J.-M. Density Functional Theory–Broken Symmetry (DFT–BS) Methodology Applied to Electronic and Magnetic Properties of Bioinorganic Prosthetic Groups. In *Metalloproteins: Methods and Protocols*; Fontecilla-Camps, J. C., Nicolet, Y., Eds.; Humana Press: Totowa, NJ, 2014; pp 269–296. https://doi.org/10.1007/978-1-62703-794-5_15.
(133) Hobbs, D.; Kresse, G.; Hafner, J. Fully Unconstrained Noncollinear Magnetism within the Projector Augmented-Wave Method. *Phys. Rev. B* **2000**, *62* (17), 11556–11570. https://doi.org/10.1103/PhysRevB.62.11556.
(134) Boettger, J. C. Approximate Two-Electron Spin-Orbit Coupling Term for Density-Functional-Theory DFT Calculations Using the Douglas-Kroll-Hess Transformation. *Phys. Rev. B* **2000**, *62* (12), 7809–7815. https://doi.org/10.1103/PhysRevB.62.7809.

(135) Jha, G.; Heine, T. DFTB Parameters for the Periodic Table: Part III, Spin-Orbit Coupling. *J. Chem. Theory Comput.* **2022**, *18* (7), 4472–4481. https://doi.org/10.1021/acs.jctc.2c00376.
(136) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Petersson, G.; Nakatsuji, H.; others. Gaussian Development Version J. 16. Gaussian. *Inc Wallingford CT* **2020**.
(137) Seeger, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XVIII. Constraints and Stability in Hartree–Fock Theory. *J. Chem. Phys.* **1977**, *66* (7), 3045–3050. https://doi.org/10.1063/1.434318.
(138) Rock, P. A. The Standard Oxidation Potential of the Ferrocyanide-Ferricyanide Electrode at 25° and the Entropy of Ferrocyanide Ion. *J. Phys. Chem.* **1966**, *70* (2), 576–580. https://doi.org/10.1021/j100874a042.
(139) Ochterski, J. W. Thermochemistry in Gaussian. **2000**.
(140) Yaji, K.; Harasawa, A.; Kuroda, K.; Li, R.; Yan, B.; Komori, F.; Shin, S. Rashba Spin Splitting of L-Gap Surface States on Ag(111) and Cu(111). *Phys. Rev. B* **2018**, *98* (4), 041404. https://doi.org/10.1103/PhysRevB.98.041404.
(141) Lodesani, A.; Anders, G.; Bougas, L.; Lins, T.; Budker, D.; Fierlinger, P.; Aiello, C. D. Weak Magnetic Field Effects in Biology Are Measurable—Accelerated Xenopus Embryogenesis in the Absence of the Geomagnetic Field. bioRxiv October 16, 2024, p 2024.10.10.617626. https://doi.org/10.1101/2024.10.10.617626.
(142) Aiello, C. D.; Abendroth, J. M.; Abbas, M.; Afanasev, A.; Agarwal, S.; Banerjee, A. S.; Beratan, D. N.; Belling, J. N.; Berche, B.; Botana, A.; Caram, J. R.; Celardo, G. L.; Cuniberti, G.; Garcia-Etxarri, A.; Dianat, A.; Diez-Perez, I.; Guo, Y.; Gutierrez, R.; Herrmann, C.; Hihath, J.; Kale, S.; Kurian, P.; Lai, Y.-C.; Liu, T.; Lopez, A.; Medina, E.; Mujica, V.; Naaman, R.; Noormandipour, M.; Palma, J. L.; Paltiel, Y.; Petuskey, W.; Ribeiro-Silva, J. C.; Saenz, J. J.; Santos, E. J. G.; Solyanik-Gorgone, M.; Sorger, V. J.; Stemer, D. M.; Ugalde, J. M.; Valdes-Curiel, A.; Varela, S.; Waldeck, D. H.; Wasielewski, M. R.; Weiss, P. S.; Zacharias, H.; Wang, Q. H. A Chirality-Based Quantum Leap. *ACS Nano* **2022**, *16* (4), 4989–5035. https://doi.org/10.1021/acsnano.1c01347.
(143) Bangruwa, N.; Suryansh; Peralta, M.; Gutierrez, R.; Cuniberti, G.; Mishra, D. Sequence-Controlled Chiral Induced Spin Selectivity Effect in Ds-DNA. *J. Chem. Phys.* **2023**, *159* (4), 044702. https://doi.org/10.1063/5.0157931.
(144) Guo, A.-M.; Sun, Q. Spin-Selective Transport of Electrons in DNA Double Helix. *Phys. Rev. Lett.* **2012**, *108* (21), 218102. https://doi.org/10.1103/PhysRevLett.108.218102.
(145) Naaman, R.; Paltiel, Y.; Waldeck, D. H. Chiral Molecules and the Electron Spin. *Nat. Rev. Chem.* **2019**, *3* (4), 250–260. https://doi.org/10.1038/s41570-019-0087-1.
(146) Dednam, W.; Lombardi, E. B.; Zotti, L. A.; García-Blázquez, M. A.; Palacios, J. J. Magnetoconductance from Spin-Charge Inter-Conversion in Two-Terminal Molecular Nanojunctions with Strong Spin-Orbit Coupling Electrodes. *J. Phys. Conf. Ser.* **2024**, *2701* (1), 012095. https://doi.org/10.1088/1742-6596/2701/1/012095.

(147) Gu, Y.; Guberman-Pfeffer, M. J.; Srikanth, V.; Shen, C.; Giska, F.; Gupta, K.; Londer, Y.; Samatey, F. A.; Batista, V. S.; Malvankar, N. S. Structure of Geobacter Cytochrome OmcZ Identifies Mechanism of Nanowire Assembly and Conductivity. *Nat. Microbiol.* **2023**, *8* (2), 284–298. https://doi.org/10.1038/s41564-022-01315-5.

(148) Pontes, R. B.; Rocha, A. R.; Sanvito, S.; Fazzio, A.; da Silva, A. J. R. Ab Initio Calculations of Structural Evolution and Conductance of Benzene-1,4-Dithiol on Gold Leads. *ACS Nano* **2011**, *5* (2), 795–804. https://doi.org/10.1021/nn101628w.

(149) Damle, P.; Ghosh, A. W.; Datta, S. First-Principles Analysis of Molecular Conduction Using Quantum Chemistry Software. *Chem. Phys.* **2002**, *281* (2), 171–187. https://doi.org/10.1016/S0301-0104(02)00496-2.

(150) Yang, G.; Wu, H.; Wei, J.; Zheng, J.; Chen, Z.; Liu, J.; Shi, J.; Yang, Y.; Hong, W. Quantum Interference Effect in the Charge Transport through Single-Molecule Benzene Dithiol Junction at Room Temperature: An Experimental Investigation. *Chin. Chem. Lett.* **2018**, *29* (1), 147–150. https://doi.org/10.1016/j.cclet.2017.06.015.

(151) Yeganeh, S.; Ratner, M. A.; Galperin, M.; Nitzan, A. Transport in State Space: Voltage-Dependent Conductance Calculations of Benzene-1,4-Dithiol. *Nano Lett.* **2009**, *9* (5), 1770–1774. https://doi.org/10.1021/nl803635t.

(152) Kim, Y.; Pietsch, T.; Erbe, A.; Belzig, W.; Scheer, E. Benzenedithiol: A Broad-Range Single-Channel Molecular Conductor. *Nano Lett.* **2011**, *11* (9), 3734–3738. https://doi.org/10.1021/nl201777m.

(153) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585* (7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2.

(154) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.

(155) Livernois, W. GauNEGF, 2025. https://doi.org/10.5281/zenodo.15178666.

(156) Palacios, J. J.; Pérez-Jiménez, A. J.; Louis, E.; SanFabián, E.; Vergés, J. A. First-Principles Approach to Electrical Transport in Atomic-Scale Nanostructures. *Phys. Rev. B* **2002**, *66* (3), 035322. https://doi.org/10.1103/PhysRevB.66.035322.

(157) Papior, N.; Lorente, N.; Frederiksen, T.; García, A.; Brandbyge, M. Improvements on Non-Equilibrium and Transport Green Function Techniques: The next-Generation Transiesta. *Comput. Phys. Commun.* **2017**, *212*, 8–24. https://doi.org/10.1016/j.cpc.2016.09.022.

(158) Taylor, J.; Guo, H.; Wang, J. *Ab Initio* Modeling of Quantum Transport Properties of Molecular Electronic Devices. *Phys. Rev. B* **2001**, *63* (24), 245407. https://doi.org/10.1103/PhysRevB.63.245407.
(159) Damle, P. S. Nanoscale Device Modeling: From MOSFETs to Molecules. *Theses Diss. Available ProQuest* **2003**, 1–131.
(160) Jacob, D.; Palacios, J. J. Critical Comparison of Electrode Models in Density Functional Theory Based Quantum Transport Calculations. *J. Chem. Phys.* **2011**, *134* (4), 044118. https://doi.org/10.1063/1.3526044.
(161) Pulay, P. Convergence Acceleration of Iterative Sequences. the Case of Scf Iteration. *Chem. Phys. Lett.* **1980**, *73* (2), 393–398. https://doi.org/10.1016/0009-2614(80)80396-4.
(162) Slater, J. C.; Koster, G. F. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **1954**, *94* (6), 1498–1524. https://doi.org/10.1103/PhysRev.94.1498.
(163) Papaconstantopoulos, D. A. The 5d Transition Metals. In *Handbook of the Band Structure of Elemental Solids: From Z = 1 To Z = 112*; Papaconstantopoulos, D. A., Ed.; Springer US: Boston, MA, 2015; pp 251–303. https://doi.org/10.1007/978-1-4419-8264-3_6.
(164) Ross, R. B.; Powers, J. M.; Atashroo, T.; Ermler, W. C.; LaJohn, L. A.; Christiansen, P. A. Ab Initio Relativistic Effective Potentials with Spin–Orbit Operators. IV. Cs through Rn. *J. Chem. Phys.* **1990**, *93* (9), 6654–6670. https://doi.org/10.1063/1.458934.
(165) Pitzer, R. M.; Winter, N. W. Spin-orbit (Core) and Core Potential Integrals. *Int. J. Quantum Chem.* **1991**, *40* (6), 773–780. https://doi.org/10.1002/qua.560400606.
(166) Fernández-Seivane, L.; Oliveira, M. A.; Sanvito, S.; Ferrer, J. On-Site Approximation for Spin–Orbit Coupling in Linear Combination of Atomic Orbitals Density Functional Methods. *J. Phys. Condens. Matter* **2006**, *18* (34), 7999–8013. https://doi.org/10.1088/0953-8984/18/34/012.
(167) Figgen, D.; Rauhut, G.; Dolg, M.; Stoll, H. Energy-Consistent Pseudopotentials for Group 11 and 12 Atoms: Adjustment to Multi-Configuration Dirac–Hartree–Fock Data. *Chem. Phys.* **2005**, *311* (1), 227–244. https://doi.org/10.1016/j.chemphys.2004.10.005.
(168) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24* (6), 669–681. https://doi.org/10.1002/jcc.10189.
(169) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals. *J. Chem. Phys.* **2006**, *125* (22), 224106. https://doi.org/10.1063/1.2404663.
(170) Wang, Y.; Demir, B.; Mohammad, H.; Oren, E. E.; Anantram, M. P. Computational Study of the Role of Counterions and Solvent Dielectric in Determining the Conductance of B-DNA. *Phys. Rev. E* **2023**, *107* (4), 044404. https://doi.org/10.1103/PhysRevE.107.044404.
(171) Ross, R. B.; Powers, J. M.; Atashroo, T.; Ermler, W. C.; LaJohn, L. A.; Christiansen, P. A. Ab Initio Relativistic Effective Potentials with Spin–Orbit Operators. IV. Cs through Rn. *J. Chem. Phys.* **1990**, *93* (9), 6654–6670. https://doi.org/10.1063/1.458934.

(172) Mtangi, W.; Kiran, V.; Fontanesi, C.; Naaman, R. Role of the Electron Spin Polarization in Water Splitting. *J. Phys. Chem. Lett.* **2015**, *6* (24), 4916–4922. https://doi.org/10.1021/acs.jpclett.5b02419.
(173) Zainal, B. S.; Ker, P. J.; Mohamed, H.; Ong, H. C.; Fattah, I. M. R.; Rahman, S. M. A.; Nghiem, L. D.; Mahlia, T. M. I. Recent Advancement and Assessment of Green Hydrogen Production Technologies. *Renew. Sustain. Energy Rev.* **2024**, *189*, 113941. https://doi.org/10.1016/j.rser.2023.113941.
(174) Okuta, R.; Unno, Y.; Nishino, D.; Hido, S.; Loomis, C. CuPy: A NumPy-Compatible Library for NVIDIA GPU Calculations. In *Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Thirty-first Annual Conference on Neural Information Processing Systems (NIPS)*; 2017.
(175) Choudhary, K.; Garrity, K. F.; Reid, A. C. E.; DeCost, B.; Biacchi, A. J.; Hight Walker, A. R.; Trautt, Z.; Hattrick-Simpers, J.; Kusne, A. G.; Centrone, A.; Davydov, A.; Jiang, J.; Pachter, R.; Cheon, G.; Reed, E.; Agrawal, A.; Qian, X.; Sharma, V.; Zhuang, H.; Kalinin, S. V.; Sumpter, B. G.; Pilania, G.; Acar, P.; Mandal, S.; Haule, K.; Vanderbilt, D.; Rabe, K.; Tavazza, F. The Joint Automated Repository for Various Integrated Simulations (JARVIS) for Data-Driven Materials Design. *Npj Comput. Mater.* **2020**, *6* (1), 1–13. https://doi.org/10.1038/s41524-020-00440-1.
(176) Vogel, E. Technology and Metrology of New Electronic Materials and Devices. *Nat. Nanotechnol.* **2007**, *2* (1), 25–32. https://doi.org/10.1038/nnano.2006.142.
(177) Xiao, L.; Hu, X.; Zhou, Z.; Xie, X.; Huang, S.; Ji, M.; Xu, A.; Tian, Y. Diverse Applications of DNA Origami as a Cross-Disciplinary Tool. *Nanoscale* **2025**. https://doi.org/10.1039/D4NR04490H.
(178) Livernois, W.; Cao, P. (Simon); Saha, S.; Ding, Q.; Gopinath, A.; Anantram, M. P. Ion Detection in a DNA Nanopore FET Device. *Nanotechnology* **2024**, *35* (32), 325202. https://doi.org/10.1088/1361-6528/ad460b.
(179) De, A.; Lu, B.; Ohayon, Y. P.; Woloszyn, K.; Livernois, W.; Yang, C.; Perren, L.; Mao, C.; Botana, A. S.; Hihath, J.; Canary, J. W.; Sha, R.; Anantram, M. P.; Vecchioni, S. Transmetalation for DNA-Based Molecular Electronics. ChemRxiv January 2, 2025. https://doi.org/10.26434/chemrxiv-2025-nkbgv.