Application of quantum computing to biochemical systems: A look to the future

Hai-Ping Cheng [1], **Erik Deumens** [1], **J. K. Freericks** [2,*], **Chenglong Li** [3], **and Beverly Sanders** [4]

[1]*Quantum Theory Project, Department of Physics, University of Florida, Gainesville, Florida, USA*
[2]*Department of Physics, Georgetown University, 37th and O Sts. NW, Washington DC 20057, USA*
[3]*Department of Medicinal Chemistry, University of Florida, Gainesville, Florida, USA*
[4]*Department of Computer Science and Engineering, University of Florida, Gainesville, Florida, USA*

Correspondence*:
J. K. Freericks
james.freericks@georgetown.edu

ABSTRACT

Chemistry has been viewed as one of the most fruitful near-term applications to science of quantum computing. Recent work in transitioning classical algorithms to a quantum computer has led to great strides in improving quantum algorithms and illustrating their quantum advantage. Much less effort has been placed on how one finishes these calculations by using the results from the quantum computer (on the active region of the molecule) and embeds them back into the remainder of the molecule in order to determine the properties of the entire molecule. Such strategies are critical if one wants to expand the focus to biochemical molecules that contain active regions that cannot be properly explained with classical algorithms on classical computers. While we do not solve this problem here, we provide an overview of where the field is going to enable such problems to be tackled in the future.

Keywords: computational molecular biology, biochemistry, quantum computing, hybrid quantum-classical algorithms

1 INTRODUCTION

Biochemical systems are essential for carrying out biological functions, and their actions span extreme time and length scales. These systems consist of proteins, DNAs, RNAs, carbohydrates, or lipids (either individually or in combination) with small molecule ligands and/or with ions in aqueous or membrane environments. The functional processes can be either covalent or not, such as molecular recognition; or a combination of both, such as an enzymatic cycle. In order to understand these elementary processes, together with experimental approaches, various computational methods have been developed at the electronic, the atomic, and more coarse-grained levels over the decades. However, full quantum calculations are intractable due to the large molecule sizes and the high demands for accuracy required for chemical applications. The solution may lie in quantum computing: as Feynman once said, "…Nature isn't classical…, …if you want to make a simulation of Nature, you'd better make it quantum mechanical…" [Feynman (1982)]. As a matter of fact, from the remarkable speed of enzyme-catalyzed reactions to the workings of the human brain, numerous biological puzzles are now being explored for evidence of quantum effects. Well-known examples include photosynthesis, nitrogen fixation, magnetoreception, olfaction, neuronal signal processing, and so on. There have even been early attempts to develop quantum computing algorithms specifically for photosynthesis and nitrogen fixation [Reiher et al. (2017)].

2 EXAMPLES

We present three representative biochemical systems in which quantum computing could be superior for accurate simulation due to the need to deal with complicated electron correlation. The first two are open

shell transition-metal and conjugated pi-electron strongly correlated systems; the last one displays extreme non-covalent intermolecular binding involving a large number of atoms.

2.1 Transition-metal-ion-containing enzyme: histone demethylase

Histone demethylases are enzymes that remove methyl (-CH3) groups from histones. The demethylase proteins alter transcriptional regulation of the genome by controlling the methylation levels that occur on DNA and/or histones and, in turn, regulate the chromatin states at specific gene loci within organisms. The big demethylase family has KDM1-6 classes [Pedersen and Helin (2010)]. Defined by their mechanisms, two main classes of histone demethylases exist: a flavin adenine dinucleotide (FAD)-dependent amine oxidase, and an Fe(II) and α-ketoglutarate-dependent hydroxylase. Both operate by hydroxylation of a methyl group followed by dissociation of formaldehyde. By studying various demethylation details, improvements are possible in the understanding of how "histone code" is employed for gene on/off switching.

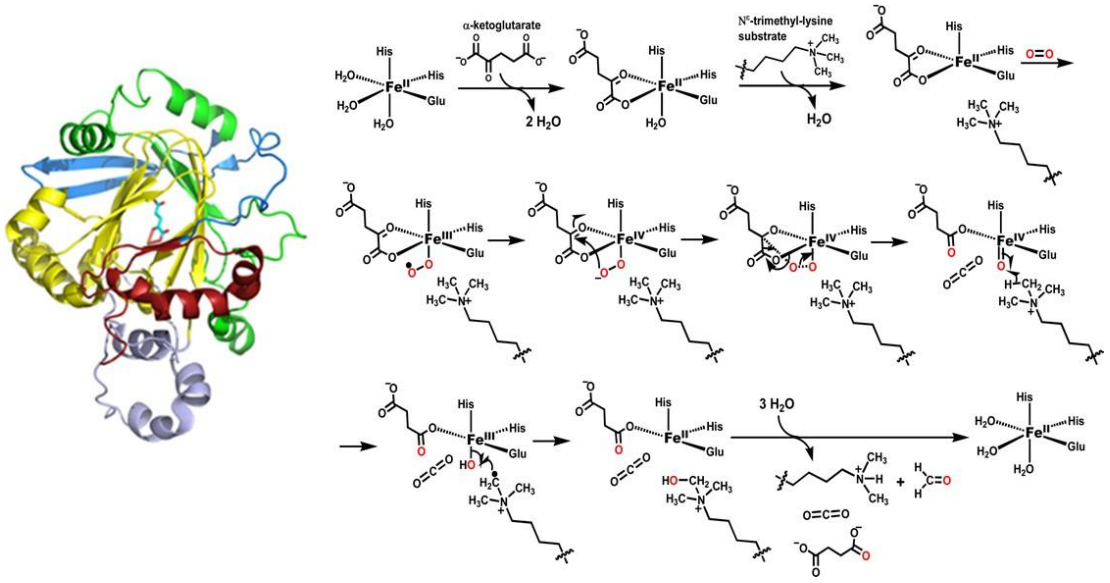

Figure 1. Structure of JmJD2A. Some domains from above are highlighted: JmJ(N-terminus, red; C-terminus, yellow), Zinc finger domain (light purple), Beta-hairpin (light blue), and mixed domain linker (green). The ball-and-sticks are Fe(II) and alpha-ketoglutarate cofactors. The enzymatic reactions involve both iron redox and oxygen radical, an ideal case where understanding can be improved with quantum computing.

Figure 1 is an illustration of the JmJD2A topology, active site, and proposed catalytic mechanism which involve both transition metal ions and reaction radicals [Chen (2006); Ng (2007)]. These are cases of Born-Oppenhiemer breakdown. During the catalytic cycle, the iron metal ion has three charge states: +2, +3, and +4, and two spin states: 0 and 1/2. Oxygen has three spin states: 0, 1/2 and 1. There are at least nine catalytic steps. Considering only direct contact catalytic amino acid residues, oxygen, trimethylated quaternary amine from lysine substrate, and of course catalytic Fe ion, 151 electrons and 121 spatial orbitals must be involved to achieve accurate electronic structure and related energy calculations.

2.2 Non-metal-ion-containing enzyme: Telomerase

At each end of a chromosome, there is a region of repetitive nucleotide sequences called a telomere which protects the chromosome from deterioration or fusion with neighboring chromosomes. For vertebrates, the sequence of nucleotides in telomeres is AGGGTT [Harvey (2014)]. The complementary DNA strand

is TCCCAA, which also has a single-stranded TTAGGG overhang. [Witzany (2008)] This sequence of TTAGGG is repeated approximately 2,500 times in humans [Sadava (2011)]. During chromosome replication, the enzymes that duplicate DNA cannot continue their duplication all the way to the end of a chromosome, so after each cell division, the telomere gets shorter. They are replenished by an enzyme, as shown in figure 2, telomerase reverse transcriptase (TERT), [Cohen S (2007)] which is the catalytic subunit of telomerase (Shay, Wright, 2019, https://www.utsouthwestern.edu/labs/shay/research/facts-about-telomeres-telomerase.html)

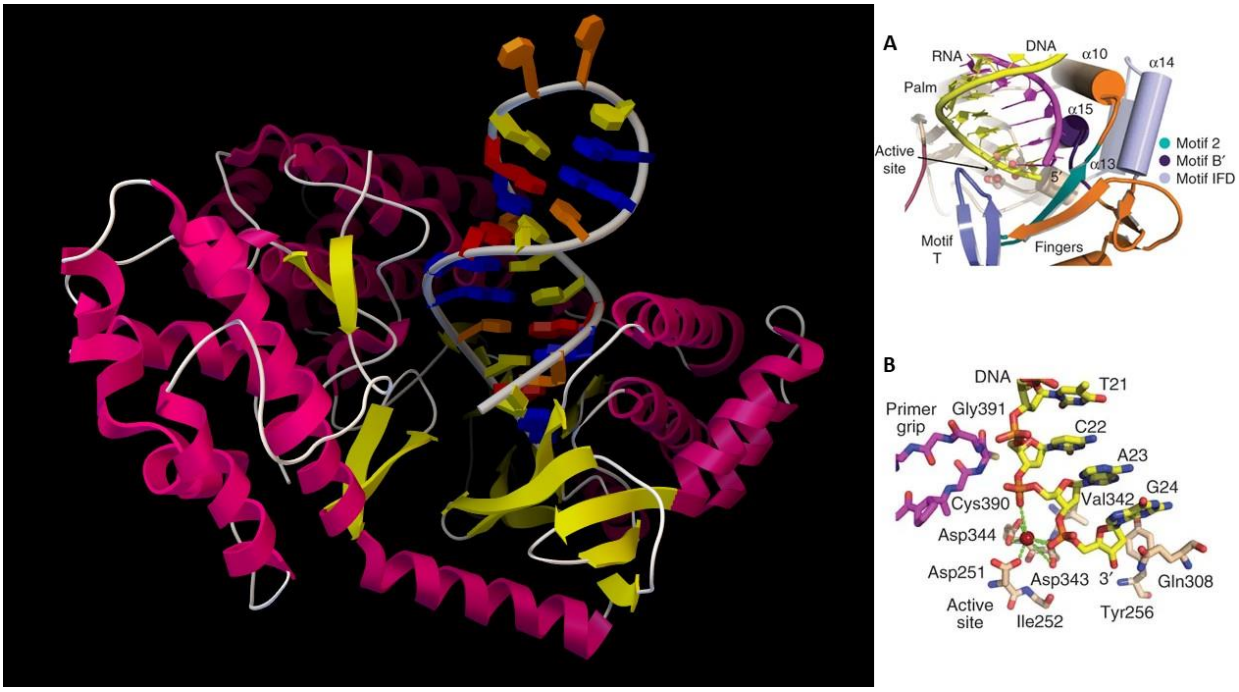

Figure 2. Left: TERT with hybrid RNA/DNA bound; Right: A. cartoon representation of the active site; B. detailed active site residues and DNA substrate.

Telomerase is active in normal stem cells and most cancer cells, but is normally absent from, or at very low levels in, most somatic cells. The study of telomerase is thus of tremendous significance for stem cell maintenance, aging and cancer. The active telomerase is a homodimer, each monomer having telomerase reverse transcriptase (TERT), telomerase RNA, and dyskerin [Mitchell (2010)]. Currently, there are several TERT crystal structures available; computational simulation of TERT telomere elongation is important and realistic. Snapshots of the molecular processes can be constructed and quantum computing can be used to simulate the catalytic active centers in order to better understand how these systems work, especially base fidelity preservation during the extension process. Due to its nucleobase pairing and reaction processivity, this is a case where quantum computing can make a large impact on molecular recognition.

2.3 Molecular recognition: Biotin-(Strept)Avidin binding

The origin of strong non-covalent reversible binding of small molecule biotin to proteins avidin (Ka 10E15 M-1) and streptavidin (Ka 10E13 M-1) remains a mystery even though many studies have attempted to resolve it since this is a classic molecular recognition issue [DeChancie (2007)]. As seen from figure 3, the beta-barrel shaped avidin binds the biotin ligand with van der Waals, electrostatic, hydrogen bonding and pi-electron polarization forces; this results in a free energy of binding around -20 Kcal/mol, almost at a quasichemical bonding level.

Understanding this binding thermodynamics at the molecular level holds fundamental importance theoretically and offers key insights for molecular design. There is no doubt that in desolvation, conformational and vibrational entropy play an important role. However, the key issue here is to understand biotin-avidin intermolecular interaction, which rests on computing accurate non-bonded interaction

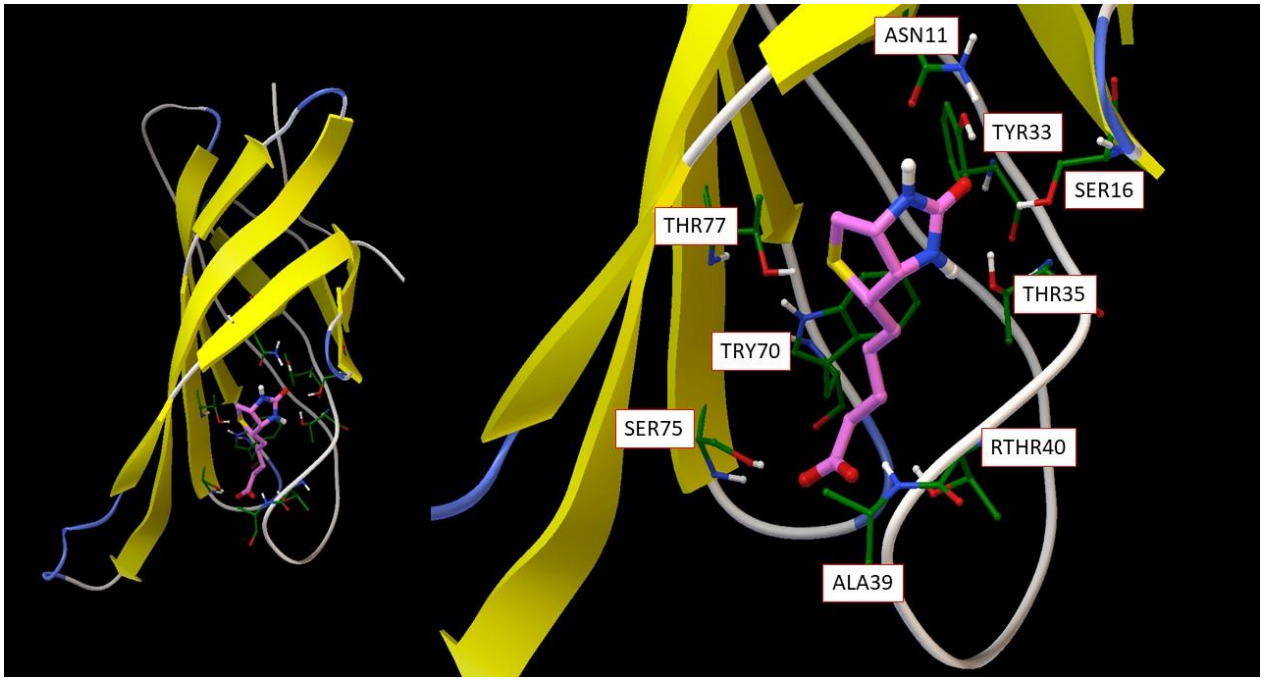

Figure 3. Left: overall avidin protein structure with biotin binding; Right: detailed biotin interaction amino acid residues from avidin.

energies. The biotin molecule (as the ligand) has 89 valence electrons and 79 frontier orbitals. By only considering direct contacting atoms from the binding amino acid residues, the active part of the molecule has a total of 379 electrons and 358 spatial orbitals. The spin state is $S = 0$ and charge state is -1. It should be obvious that such a subsystem is too large to be tackled with standard chemical methodologies for strongly correlated molecules, so quantum computing is the only option for a complete theoretical analysis.

3 EMBEDDING STRATEGIES

All embedding strategies available now require hybrid quantum-classical algorithms because large molecules will remain too large to be simulated entirely on a quantum computer for many years in the future. In addition, computational times may also be too long for such large systems. Nevertheless, by developing appropriate schemes that embed calculations involving the strongly correlated parts of the molecule into the less expensive calculations involving the rest of the molecule, one can break up the tasks into those run on a quantum computer and those run on classical machines. This section describes such embedding approaches.

The molecular systems of interest in biological processes are complex as well as geometrically extensive. It is also known that some processes depend crucially on small differences in structure and their associated energy differences. The timescales of processes may also span several orders of magnitude, from femtoseconds for molecular vibrational changes to milliseconds for some electron transfer processes and conformation changes. The challenge is that highly accurate calculations are needed for these extreme systems.

In addition, biological processes happen at finite temperature in a liquid environment, as opposed to many chemical processes that can be understood by studying the gas phase or materials structures that are often analyzed in isolation and at absolute zero. This means that a statistical description is needed to describe the full process and to obtain accurate reaction rates. Entropy and free energy play a crucial role in this.

Hence, biological molecules appear to be an ideal application for the promised power of quantum computing. However, with noisy intermediate-scale quantum computers (NISQ), such calculations are

currently out of reach. Even when more fully fault-tolerant quantum computers become available, it is likely that the complete statistical quantum description of realistic biomolecular systems and processes will require embedding strategies within a classical/quantum hybrid approach.

3.1 Multiscale simulation and modeling

The method of multiscale simulation and modeling may offer some insights on how to tackle a complex problem like this. Here the challenge is to describe the processes that are visible at the macroscopic level but are fully determined by the details at some microscopic level. A great example is the formation and propagation of cracks in materials [Budarapu et al. (2014); Gao and Klein (1998); Liu et al. (2004); Rountree et al. (2002); Rudd and Broughton (2000); Talebi et al. (2014)]. The macroscopic description is of interest, but the continuum models that are effective and affordable at that scale cannot describe the basic bond breaking process that lies at the foundation of the crack formation. Nor can the molecular dynamics methods describe this process. Thus the continuum model, the molecular dynamics model, and the quantum model must be coupled and scale-bridging techniques used to connect them in a way that accurately preserves the physics [Hoekstra et al. (2014)].

Similarly, a biomolecule can be divided into three regions: a classical region where interatomic interaction can be treated with classical force fields using standard methods [Amber http://ambermd.org/, CHARMM https://www.charmm.org/charmm/, LAMMPS https://lammps.sandia.gov/, etc]; a quantum region where mean-field approximations are sufficient; and a strongly correlated region where high-level methods that treat quantum entanglement are needed, that is, techniques beyond density functional theory (DFT). To begin, we need to examine the feasibility of different approaches in the second quantization formalism. First, DFT with a constrained random phase approximation (cRPA) can be used to obtain the single-particle Hamiltonian H_0 and the two-particle U and J interaction matrices. Second, quantum chemistry methods (e.g. the coupled-cluster method) can be used to obtain the Fock matrix and the exchange and Coulomb terms. In both of these approaches, one uses a Hamiltonian downfolding technique to reduce the dimensionality of the Hilbert space. This practice was originally developed in solid-state physics to minimize the number of energy bands in a generalized Hubbard model. In solid state-physics, these methods have been applied to transition-metal oxides [Zhang et al. (2019b)] (MnO, FeO, CoO and NiO) as well as molecular magnetic systems for studies of magnetic couplings. Below we provide some details of the multi-scale simulation scheme.

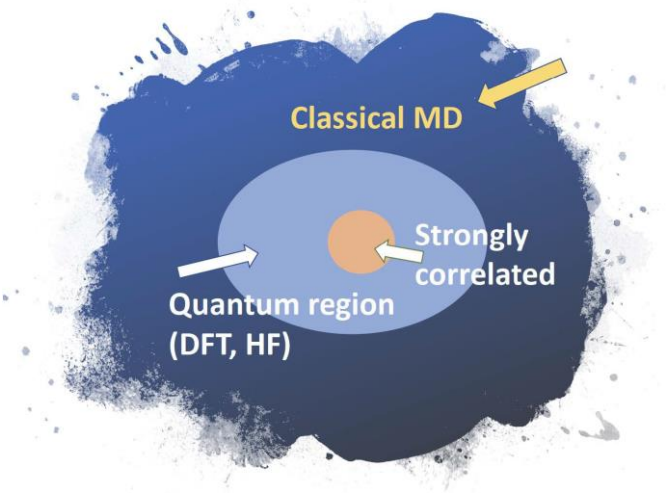

Figure 4. Schematic of a hybrid simulation framework for molecules that employ a hierarchical embedding strategy.

3.1.1 Hybrid quantum-classical molecular dynamics

We illustrate how hierarchical methods are used in a study of water-silica surface interaction. The system is divided into two regions, the quantum and the classical. Here, the quantum region is described by effective mean-field methods, while the classical region is described by molecular dynamics using effective force laws. The two regions must be coupled together across their boundary. Various methods exist for

the embedding of the quantum region (the light blue region in Figure 4) inside the classical region. In earlier work [Du et al. (2004)], a quantum region described by DFT is embedded in a classical matrix as shown in Figure 5. This figure depicts a Si-O bond-breaking process on the silica surface. According to a free cluster model [Walsh et al. (2000)], the calculated barrier energy of this process is E_b=0.7-1.1 eV. However, when the cluster is embedded in a surface matrix, the calculated barrier energy E_b is equal to 0.4 eV. Including quantum effects results in a substantial decrease. For a bio-molecule, the embedding is simpler than for amorphous materials as there are not as many bonds that connect the classical region and the quantum region. Techniques for this type of embedding are quite sophisticated [Bakowies and Thiel (1996); Cui et al. (2001); Friesner and Guallar (2005); Gao et al. (1998); Gao and Xia (1992); Laio et al. (2002); Vreven et al. (2003)] and should be easy to adapt to the biomolecule application.

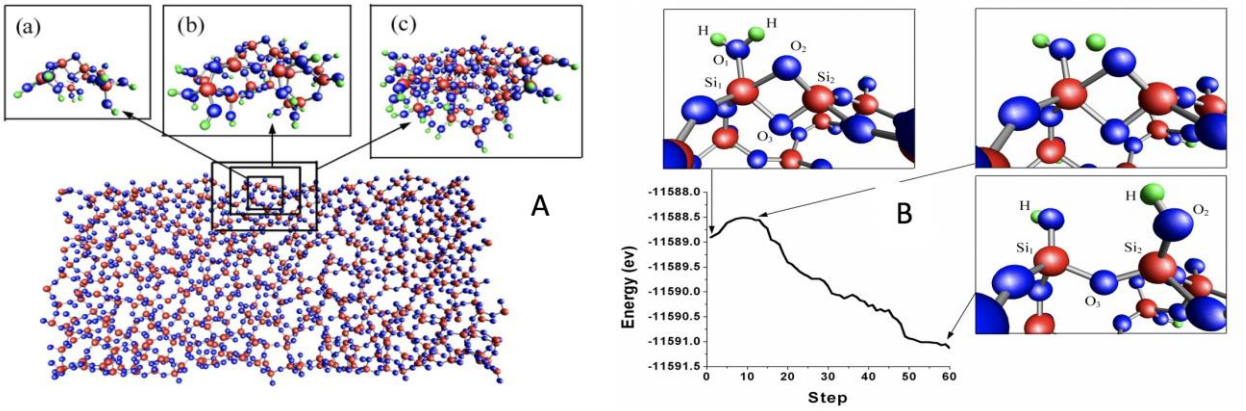

Figure 5. Sketch of the hybrid simulation framework as applied to amorphous glasses. Three sizes of quantum region were chosen (panels a-b on the left) to ensure the convergence of the reaction energy. The right picture relates energy path with water splitting process. The energy barrier is only 0.4 eV and is zero when the reaction involves two water molecules

3.1.2 From DFT to highly-correlated systems: Quantum Embedding theory for molecules and for the Hubbard model

Transition-metal molecular complexes with d and f electrons often demonstrate strong correlation effects. Active centers of many enzymes are transition-metal complexes; *e.g.*, photosystem I and II have iron-sulfur and manganese-oxide clusters as their active centers. The large number of atoms in ligands makes high-level calculation of the whole molecule impossible. In this situation, quantum embedding is necessary. Note that quantum embedding is different from the hybrid quantum-classical simulation discussed in section 3.1.1. Here, we embed a strongly correlated subspace in a single electron space. So, we need a single-particle theory for the whole molecule and a many-body theory for the correlated subspace, which is a small but functional part of the molecule. One embedding scheme utilizes density functional theory (DFT) as the single-particle theory and does the embedding via dynamical mean field theory (DMFT) [Georges et al. (1996)]. In DMFT, the correlated subspace is referred as an impurity and the impurity problem is solved by an impurity solver. We can use unitary coupled-cluster theory, which can be calculated in quantum computers, as the impurity solver. Using a wave-function based impurity solver, we calculate the ground state $|0\rangle$ with energy E_0, and a few of low-lying excited states. Then, the impurity Green's function can be constructed as:

$$G_{imp}(i\omega_n) = \left\langle 0 \left| c_\sigma \frac{1}{i\omega_n + E_0 - \hat{H}} c_\sigma^\dagger \right| 0 \right\rangle + \left\langle 0 \left| c_\sigma^\dagger \frac{1}{i\omega_n - E_0 + \hat{H}} c_\sigma \right| 0 \right\rangle \quad (1)$$

with non-interacting Green's function defined by

$$G^0(i\omega_n) = (i\omega_n - \hat{H})^{-1} \quad (2)$$

we can extract self-energy from the impurity

$$\Sigma(i\omega_n) = \frac{1}{G^0(i\omega_n)} - \frac{1}{G_{imp}(i\omega_n)} \quad (3)$$

The impurity self-energy is approximated as the self-energy of the whole molecule in DMFT, and we can use it to obtain the molecular interacting Green's function, which in turn will be used to calculate physical properties of the molecule. This DFT+DMFT scheme has already had a lot of success in strongly correlated materials, but applications in molecular systems are still rare. We plan to use the DFT+DMFT scheme to study transition-metal complexes. The objective of this work is to identify effects of electron correlations in molecules. DFT+DMFT is very good at capturing qualitative features of electron correlation. However, the quantitative energetics of this method is limited due to uncertainty in the interaction parameters and the double-counting issue. To solve those issue, one feasible approach is to use GW [Hedin (1965)] as the single-particle theory, instead of DFT.

One of the breakthroughs in many-body physics occurred when the problem was investigated in the limit of infinite spatial dimensions. It was then found that the spatial correlations are so suppressed that one can limit to an exact treatment of local but temporal fluctations. This technique is the dynamical mean-field theory (DMFT) which was developed for extended systems first. DMFT incorporates a self-consistently determined quantum impurity problem, similar to the Anderson impurity model, which is embedded in an effective bath. The impurity model is defined, in part, from the on-site Coulomb interaction U -matrix. One reliable way to determine these parameters from first principles is the constrained Random Phase Approximation (cRPA) method [Aryasetiawan et al. (2004, 2006)]. In the original idea, one aims to estimate the screened Coulomb interaction for selected bands of interest, that is, within a specified energy window. For this purpose, the particle-hole polarization between all possible pairs of occupied and unoccupied states is taken into account. Within the Random Phase Approximation (RPA), the particle-hole polarization is calculated via

$$P(r, r; \omega) = \sum_{i}^{\text{occ.}} \sum_{j}^{\text{unocc.}} [\psi_i^*(r)\,\psi_j(r)\,\psi_i^*(r)\,\psi_i(r)] \left(\frac{1}{\omega - \varepsilon_j + \varepsilon_i + i\delta} + \frac{1}{\omega + \varepsilon_j - \varepsilon_i - i\delta} \right) \tag{4}$$

[Petersilka et al. (1996)]. Here, ψ_i and ε_i are the eigenfunctions and eigenenergies of the single-particle Hamiltonian of the underlying DFT. The summations over i and j are restricted such that i must be an occupied state (real orbital) and j must be an unoccupied state (virtual orbital).

The bands of interest are often found near the Fermi level, and have a particular orbital character, for example, d-like for transition-metal oxides. We label the bands of interest as the d-space. If both the occupied states and the unoccupied states are within the d-space, then their polarization contributes to $P_d(r, r; \omega)$. All other pairs of occupied and unoccupied states contribute to P_r, where r stands for the rest of the bands. This naturally separates the total polarization into two parts: $P = P_d + P_r$. The P_r is the quantity related to the partially screened Coulomb interaction, given by [Aryasetiawan et al. (2004)].

$$W_r(\omega) = [1 - v\,P_r(\omega)]^{-1}\,v \tag{5}$$

Here, v is the bare Coulomb interaction. According to the Hedin equations and the GW approximation [Hedin (1965)], the total polarization P_r screens the bare Coulomb interaction, v, to give the fully screened interaction W. With the same logic, P_d screens W_r to give the fully screened interaction W. Thus, W_r is identified as the screened on-site Coulomb interaction for the d-space, that is, $U(\omega) \equiv W_r(\omega)$, which includes the screening effect from the realistic environment of the real material.

To make DFT+DMFT fully *ab initio*, the hopping parameters and the Coulomb interaction parameters should be provided from first principles (in the DFT part of the calculation). The Hubbard Hamiltonian can be written as

$$\hat{H}_{\text{Hubbard}} = \sum_{i,j} t_{ij}\,\hat{c}_i^\dagger\,\hat{c}_j + \sum_{i,\alpha,\beta,\gamma,\delta} U_{i,\alpha\beta\gamma\delta}\,\hat{c}_{i,\alpha}^\dagger\,\hat{c}_{i,\beta}^\dagger\,\hat{c}_{i,\gamma}\,\hat{c}_{i,\delta}. \tag{6}$$

The hopping matrix t_{ij} comes from the DFT eigenenergies and provides the bath Green's function in DMFT. The Coulomb interaction $U_{i,\alpha\beta\gamma\delta}$ comes from the cRPA calculation described above, which is the only fully quantum-mechanical way to obtain the Coulomb interaction parameters. With the bath Green's function and U at hand, the effective action of the impurity problem is constructed. DMFT solves for the impurity Green's function by direct numerical sampling of the Green's function $G_{ij,\sigma}(t) = -\langle T\hat{c}_{i,\sigma}(t)\,\hat{c}_{j,\sigma}^\dagger(0)\rangle$ using the continuous time quantum Monte Carlo algorithm (CT-QMC). [Gull et al. (2011); Zhang et al. (2019a)]

It has been proposed [Bauer et al. (2016)] that a quantum computer algorithm could replace the CT-QMC calculation and provide the impurity Green's function $G_{ij,\sigma}(t)$, especially in cases where the classical computation suffers from the sign problem. Such a calculation embeds the impurity solver onto the quantum computer (quantum computing task), while the remainder of the DFT+DMFT iteration is carried out on classical computers. However, because describing the bath for the impurity problem is complex, it might be fruitful to instead simply solve the many-band lattice problem directly on the quantum computer. Indeed, this latter approach is more likely to be generalizable to large molecular systems.

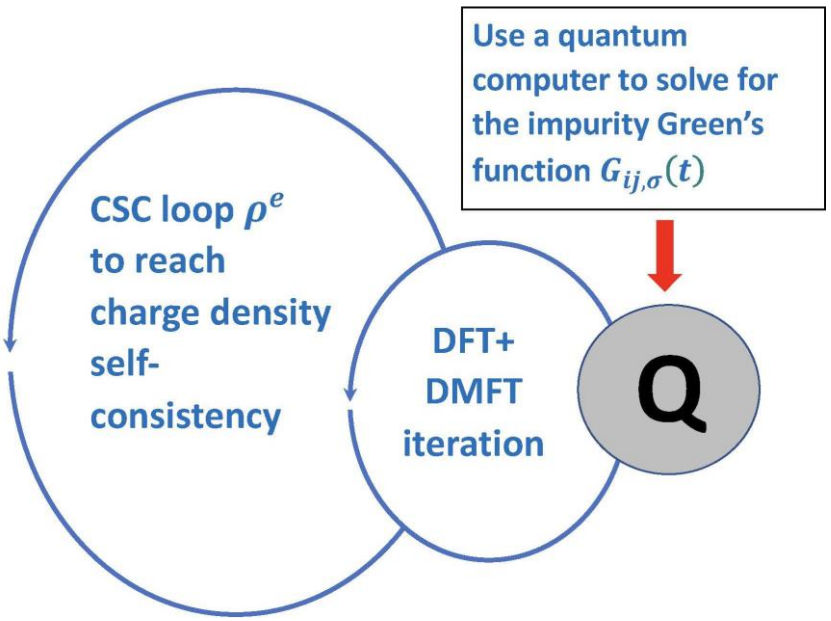

Figure 6. no caption yet

It often is important to embed the DFT+DMFT iteration into a larger loop of charge-density (ρ^e) self-consistency (CSC) (Figure 6). It is known that CSC DFT+DMFT is necessary to capture charge density re-distributions even for very simple transition metal oxides like V_2O_3 under ambient conditions [Leonov et al. (2015)]. Similarly, in molecular calculation, charge redistribution is important for any simulations involving catalysis or other reactions.

3.1.3 Coupled-cluster Hamiltonian in second quantization form

In the coupled cluster method, the Coulomb energy J is given by

$$\hat{J} = \sum_{p,q;r,s} (pq\,|\,rs)\, p^\dagger q^\dagger rs = \sum C^*_{\mu p} C^*_{\nu q} C_{\lambda r} C_{\sigma s}\, (\mu\nu\,|\,\lambda\sigma)\, p^\dagger q^\dagger rs \tag{7}$$

where

$$(\mu\nu\,|\,\lambda\sigma) = \int\int x_\mu(r_1) x_\nu(r_2) \frac{1}{r_{12}} x_\lambda(r_1) x_\sigma(r_2)\, dr_1\, dr_2 \tag{8}$$

and, p, q are molecular orbitals. In addition, $\mu, \nu, \lambda, \sigma$ are indices for basis functions. The Hamiltonian is written as

$$\hat{H} = \hat{F} p^{\dagger} q + \frac{1}{4} \quad (pq \,||\, rs)\, p^{\dagger} q^{\dagger} sr \tag{9}$$

$$(pq \,||\, rs) = (pq \,||\, rs) - (pq \,||\, sr). \tag{10}$$

The wave function becomes

$$\psi = e^{\hat{T}} |\psi_0) \tag{11}$$

where $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots.$ and $|\psi_0)$ is the Slater determinant (reference function). In principle, one can work out as many terms as computationally feasible. For quantum computers, the unitary coupled cluster method instead of the coupled cluster method method is the way to go, that is we let $\hat{T} \rightarrow \hat{T} - \hat{T}^{\dagger}$ Several groups are working on the embedding schemes.

3.2 Fragment molecular orbital method

In computational chemistry for large systems, the fragment molecular orbital (FMO) method [Gordon et al. (2012); Tanaka et al. (2014); Zahariev and Gordon (2012)] was developed to solve a similar problem: namely that the system is too large to treat as a whole with standard methods. In that case, the molecule is divided into fragments that can be chosen to (in some sense) contain atoms that interact strongly with each other, but less strongly with atoms in other fragments. Standard methods are used to obtain an accurate description of the isolated fragments. Then, other methods, also standard, are used to describe the interaction between the fragments and the effect those interactions have on the internal structure and properties of the fragments. The result converges to a solution for the complete system of all interacting fragments with controllable accuracy. In addition, the method shows linear scaling for large systems.

3.3 Quantum computing on fragments

To describe the biochemical systems, one can envision a similar approach. Now, however, instead of using different methodologies for different regions, one uses a classical computer to describe one region and a quantum processor for the region where the model can use the advantage offered by quantum computing. The self-consistency would typically be carried out on a classical computer. The approach is similar to the variational quantum eigensolver (VQE) method [McArdle et al. (2020); Grimsley et al. (2019a)] described below in section 4.1 for quantum chemistry on quantum computers: Part of the computation is performed on the classical computer, some information is extracted from that calculation and handed to a classical computer, which then performs the next part of the computation. That computation results in new values to be used for the next iteration on the quantum computer. Other algorithms have been formulated to combine the use of classical and quantum computers to solve a problem [Chen et al. (2019)].

For a hybrid description of a complex biological systems, the different parts of the computation are not only different stages in an algorithm, but also describe different spatial regions of the system. Let us call the region described on the quantum computer "primary" and the region described on the classical computer, which most often surrounds the primary region in space, the "environment." We assume that the primary region fits in the quantum computer in the sense that it has sufficient qubits to represent both the quantum state in some encoding from fermions to spins (such as Jordan-Wigner [McArdle et al. (2020)]) and all the ancillary qubits necessary to execute the chosen algorithm.

It is necessary to choose quantum-mechanical methods to represent the states of both the primary and environment regions of the biological system so that the desired accuracy for the complete system can be achieved. It is not necessary that both regions are treated with the same method, as long as the physical description is consistent. The algorithm then inevitably requires that information is exchanged between the classical and quantum computers about the state description of the respective components. The classical computer can easily provide the necessary information to the quantum computer, which usually changes the state preparation on the quantum computer. However, as with VQE, obtaining accurate information about the state of the primary region as represented on the quantum computer can be challenging if the quantum state is complicated since there is no efficient way to directly access to the entangled wavefunction stored on the qubits in the quantum processor. If a process of measurement needs to be called, then accurate

calculations may requires unacceptably large numbers of repeat runs of the program to obtain the required accuracy.

The general algorithm works as follows.

1. Specify a computational chemistry model for the environment region and initialize its state.
2. Specify a computational chemistry model for the primary region and initialize its state using the environment state parameters as needed.
3. Perform the algorithm to solve the computational chemistry model for the primary region on the quantum computer.
4. Extract the required information from the state of the primary region to perform the next iteration of the algorithm to converge the environment.
5. Perform the classical part of the algorithm for the primary region, using state information of the environment as needed.
6. Using information obtained by the classical part of the algorithm for the primary region, re-initialize the quantum computer for the quantum part of the algorithm for the primary region.
7. Repeat until the defined convergence criterion is met.

3.4 Sparse Green's function embedding schemes

One of the challenges with accurate quantum chemistry calculations is that a large percentage of the correlation energy arises from the sum of many small contributions. This arises in part, because any standard orbital basis results in fairly full single-particle and two-particle interaction matrices. Hamiltonian evolution, or even variational methods require evaluating many, many terms on the quantum computer. This leads to high-depth circuits, which become difficult to run on NISQ machines and may even be problematic on the expected fault-tolerant ones. One way around this problem is to transform the problem into a representation that is more sparse, or even to approximately force it into an extremely sparse representation. This is the idea behind the self-energy embedding theory [Tran et al. (2018)].

Starting from an inexpensive classical calculation (such as Hartree-Fock plus MP2), one computes a representation for the self-energy of the full chemical system. Next, one determines the high-frequency moments of the self-energy. For the retarded Green's function, these moments of the self-energy are often determined by parameters in the Hamiltonian itself (the constant term is exactly determined from the Hartree-Fock approximation, the zeroth moment from the interaction, the first moment involves a few two-particle correlation functions, and so on. The strategy is then to construct an extremely sparse interaction for the effective Hamiltonian. This has the full single-particle contributions, but restricts the Coulomb interaction to on-site direct or exchange interactions only. These interaction terms are chosen to require that the low-order moments are preserved in the effective model. Then, one solves for the full self-energy of the effective model and then uses the effective self-energy as the self-energy for the full system. This approach guarantees that the low-energy moments of the final description of the molecule are exactly preserved. A self-consistency scheme is employed to update the approximation, as the moments depend on some expectation values which change as the Green's functions change with each iteration of the calculation. We also note that equality of low-order moments also implies that the two Green's functions agree exactly for short times.

The way we envision using this on a quantum computer for large molecules is to apply this approach to the strongly correlated core (or strongly correlated fragments) and one ends up with a much lower depth circuit because the Hamiltonian is so much sparser. This will allow more complex systems to be simulated earlier than possible with algorithms that include the full chemical complexity. The quantum computer simulates only the sparse Hamiltonian and determines the Green's function (or self-energy), which then is sent to the classical computer for the remainder of the algorithm. Even in the future, when fault-tolerant quantum computers become available, methods like the self-energy embedding theory will remain valuable as they can significantly streamline the number of operations needed to be run on the quantum computer.

4 QUANTUM ALGORITHMS

We describe two paradigmatic algorithms for quantum chemistry on a quantum computer. The first is the variational quantum eigensolver [Peruzzo et al. (2014); McArdle et al. (2020)], which is viewed today as the best candidate for performing chemistry on so-called noisy intermediate-scale quantum (NISQ) hardware. We also discuss more accurate approaches like quantum phase estimation, which will ultimately emerge as the gold standard for quantum chemistry on a quantum computer because it can compute ground-state energies with only small systematic errors. We end this section with a discussion of how different quantum algorithms designed to compute energies can interface with different embedding strategies.

4.1 Variational quantum eigensolver

In the NISQ era, quantum computers will not be able to accurately simulate deep circuits. They also will have results that will require error/noise mitigation from state preparation and measurement errors and from infidelities in gate executions. Within this realm of quantum hardware, there is an algorithm that shows great promise—the variational quantum eigensolver algorithm [Peruzzo et al. (2014)]. This algorithm is essentially a "state preparation and then measure" algorithm leading to low-depth circuits governed primarily by the complexity of the state preparation. One starts from a single reference state (usually the Hartree-Fock state) and then creates a variational ansatz that depends on a set of variational parameters. There are many options for how to do this. The ADAPT-VQE approach [Grimsley et al. (2019b)] dynamically constructs the ansatz by iteritavely choosing operators from a pool of available operators. A unitary singles and doubles coupled cluster approach will use a Trotterized form of the unitary coupled cluster ansatz (with only singles and doubles excitations in the exponent). Factorized forms of the unitary coupled cluster approach have also been considered, but usually, these approaches are not easily restricted to certain classes of excitations and might be better thought of within an ADAPT type methodology. There also are ansatze that employ tensor-product-based wavefunctions.

However the wavefunction has been prepared on the quantum computer, we need to next measure the expectation value of the Hamiltonian to complete the calculation. The Hamiltonian is a Hermitian operator, so it cannot be evaluated directly on the quantum computer. Instead, we break it up into a sum of its mutually commuting unitary components and evaluate the expectation value of each unitary—the total expectation value is found by accumulating the total of all of the terms. As the number of orbitals increases, the number of terms in the Hamiltonian also increases and as a result, the complexity of this approach grows. To date, only quite simple molecules have been computed on available quantum hardware. The first approach was hydrogen and other simple binary and tertiary molecules [Kandala et al. (2017)]. The most complex system examined so far is H_2O in the STO-3G basis [Nam et al. (2020)].

Of course, this forms the inner loop of the full variational calculation. One must now adjust the parameters in the variational wavefunction to find the minimum of the energy. This optimization problem requires complex algorithms on classical computers because the data emerging from the quantum computer is noisy. The noise may even make it challenging to complete this class of calculation to the point where a true minimum can actually be located. The optimization problem may also suffer from "barren plateaus". It can be improved by calculating the derivative of how the energy changes when a variational parameter is changed from a matrix element measured directly on the quantum computer.

Because quantum computers have much slower clock cycles than classical computers, even with a quantum advantage for computing the results of a given measurement the quantum computations are expected to be slow. In addition, the parameters of a quantum computer often drift with time, creating additional issues associated with a changing accuracy for different expectation values over time. One may even need to correct for the drift over time or risk having data that is not accurate enough to be able to complete the outer loop of the variational cycle.

Nevertheless, this approach remains the most promising approach available for now. Until we are able to perform extensive time evolution on a quantum system, it will remain the only viable strategy for quantum chemistry on NISQ era machines.

4.2 Phase estimation

The phase estimation algorithm was invented by Kitaev in 1995 [Kitaev (1995)]. Rather than solve the traditional eigenvalue problem $H|\psi) = E|\psi)$, it elevates the problem to a unitary one, where we instead

try to determine the phase λE arising from the application of $e^{i\lambda \mathrm{H}}$ to the eigenfunction as follows:

$$e^{i\lambda \mathrm{H}}|\psi) = e^{i\lambda E}|\psi). \tag{12}$$

Obviously one needs to choose the λ parameter with care to ensure we can read the energy off without having the phase increase past 2π and we need to measure enough binary digits to have an accurate measure of the energy. In addition, since the phase estimation algorithm is a measurement of the unitary operator, it collapses onto the different eigenstates in the initial superposition, according to how much probability amplitude each eigenstate has in the initial superposition; the probability is given by the square modulus of the amplitude. Hence, one must prepare an initial state that has high overlap with the eigenstate that you want to determine the energy of. This could then involve a synergy with the variational quantum eigensolver algorithms in the following way: since the variational state is an approximation, it should have a high overlap with the true ground state, allowing it to be a good choice for the initial state that is put into the phase-estimation algorithm.

There are many benefits to the phase-estimation approach. First, it will give us an accurate estimate of the ground-state energy, with the accuracy determined by how many binary digits representing the phase are computed on the quantum computer.. Second, it projects onto the eigenstate it measures. This allows it to also be employed as a state-preparation protocol; measuring the ground-state energy also has the consequence of preparing the ground-state wavefunction directly on the quantum computer where it can then be employed for further quantum computations. For example, if the embedding strategy for self-energy embedding theory is used, one can compute the zero-temperature Green's function directly from the ground-state eigenfunction.

The challenge with phase estimation is that it requires us to be able to accurately perform time evolution. This is currently beyond the scope of available hardware and most likely will need to wait for large-scale fault-tolerant quantum computers to be available to be able to carry out such computations. Nevertheless, it is important to think through how one would work with such an algorithm now, to be ready when such hardware becomes available.

Of course, the first chemical systems put onto quantum computers are not going to be large biological molecules. But, with the development of the right algorithms for embedding, hierarchical structuring, and low-depth circuits, one might be able to advance biological science sooner than later. At the least, we should position ourselves to be able to try.

5 CHALLENGES

There are a few challenges associated with the modeling of biochemical systems. The first challenge is the accuracy required to describe the structures and processes. The standard is 1 kcal/mol or 4 kJ/mol, which in atomic units used in quantum mechanics is equal to 43.36 meV or 1 mHartree. Given that the energies of large molecules relevant in biochemistry are in the thousand Hartree range, the energies need to be calculated with a precision of 6 to 8 digits, which corresponds to a single precision IEEE floating point number on a classical computer. There are a large number of integrals with weights that are small. The contribution from each integral is small, but the sum adds up to a non-negligible contribution to the total energy. Because these numbers are obtained by a large number of floating point operations in the classical part of the computation, the minimum precision needed to perform this classical part of the calculation with controlled rounding error is 15 digits, which corresponds to the double precision floating point number on classical computers. For large molecules, relevant to biology, the integral contributions are sorted and added with small numbers first to build larger numbers that can be meaningfully added together to avoid critical round-off errors. That means that the step in the hybrid quantum-classical algorithm where values must be measured from the state in the quantum processor, these results need to be obtained with the right precision. Because the standard deviation of statistical sampling with N trials goes like $1/\sqrt{N}$, the number of measurements for a given accuracy ε is $N = \varepsilon^{-2}$. A careful analysis is needed on what precision will be needed for the various terms to get acceptable accuracy for the total energies, because the required precision directly impacts the number of measurements that will be required, with a quadratic impact on total run time. Some further research to improve the algorithm for processing of the integrals will be needed to determine the minimum precision of each wave function component that must be combined with each integral to get the correct precision of the final result.

The second challenge in biochemical structure and process analysis and design is that the systems are at some finite temperature. That means a statistical description is essential. This has been taken into account for decades in the molecular dynamics simulations [Karplus and McCammon (2002); Seabra et al. (2007); Salomon-Ferrer et al. (2013)], with the method of replica exchange being one of the leading approaches [Roe et al. (2008)].

However, chemical accuracy is not sufficient. Decades of research to design drugs, enzymes, and catalysts has not been as successful as once hoped. The likely root cause is that chemical accuracy is insufficient to distinguish the competing mechanisms from each other, especially once the proper statistics at room temperature is taken into account. If 1 kcal/mo were adequate, scientists would have made more progress in identifying new mechanisms. To make the computations really insightful, it is likely that at least one and probably two or three orders of magnitude higher accuracy is required to generate new insights into drug, enzyme, and catalyst activities and reaction mechanisms.

These considerations make it clear that biochemical structures and processes are a fertile ground of problems to use, and demonstrate, advantage of quantum computing over classical computing. It also shows the road to success will be difficult. But it promises to be wonderful journey!

6 CONCLUSION

This short review leaves us hopeful, but with many unanswered questions. It is clear that there are significant challenges that must be met before we can reap the benefits of quantum computers for biochemical applications. Nevertheless, due to the complexity involved in properly partitioning the sub units of these problems and then combining the results together, we need to start now to properly plan for how this will work. We can bring in ideas from a number of different areas where similar "divide and conquer" approaches have been tried and successfully completed. But the strategies that employ quantum co-processors to handle the most difficult parts of the calculations need to be properly thought out and structured so we can make rapid advances once the hardware is available. We did not map out a complete plan for how one can proceed. Instead, we described the different strategies that need to work together to achieve this goal. We are looking forward to seeing how this will work as everything comes together and quantum computation has important and significant impacts into biochemistry.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

Li wrote the biochemistry. Cheng wrote the multi-scale simulation. Deumens wrote the fragment method and its extension to quantum-classical hybrid application and the challenges. Freericks wrote algorithms. All authors worked on the overall structure of the paper.

FUNDING

The collaboration was supported by the National Science Foundation under grant number OMA-1936853. J.K.F. was also supported by the National Science Foundation under grant number CHE-1836497 and the McDevitt bequest at Georgetown. H.-P. Cheng was also supported by DOE/BES DE-FG02-02ER45995

ACKNOWLEDGMENTS

The authors thank the participants of the Sanibel workshop in quantum computing held February 21-22, 2020 and the students of the course CIS6930 on quantum information science for numerous discussions.

The datasets [GENERATED/ANALYZED] for this study can be found in the [NAME OF REPOSITORY] [LINK].

REFERENCES

Aryasetiawan, F., Imada, M., Georges, A., Kotliar, G., Biermann, S., and Lichtenstein, A. I. (2004). Frequency-dependent local interactions and low-energy effective models from electronic structure calculations. *Phys. Rev. B* 70, 195104. doi:10.1103/PhysRevB.70.195104

Aryasetiawan, F., Karlsson, K., Jepsen, O., and Schönberger, U. (2006). Calculations of hubbard *u* from first-principles. *Phys. Rev. B* 74, 125106. doi:10.1103/PhysRevB.74.125106

Bakowies, D. and Thiel, W. (1996). Hybrid models for combined quantum mechanical and molecular mechanical approaches. *Journal of Physical Chemistry* 100, 10580–10594. doi:10.1021/jp9536514

Bauer, B., Wecker, D., Millis, A. J., Hastings, M. B., and Troyer, M. (2016). Hybrid quantum-classical approach to correlated materials. *Phys. Rev. X* 6, 031045. doi:10.1103/PhysRevX.6.031045

Budarapu, P. R., Gracie, R., Yang, S. W., Zhuang, X. Y., and Rabczuk, T. (2014). Efficient coarse graining in multiscale modeling of fracture. *Theoretical and Applied Fracture Mechanics* 69, 126–143. doi:10.1016/j.tafmec.2013.12.004

Chen, C.-c., Shiau, S.-Y., Wu, M.-F., and Wu, Y.-R. (2019). Hybrid classical-quantum linear solver using noisy intermediate-scale quantum machines. *Scientific Reports* 9, 16251 (12 pages). doi:10.1038/s41598-019-52275-6

Chen, Z. J. W. J. H. X. D. F. K. T. S. M. M. Q. P. C. D. S. H. J. H. K. S. Y. Z. G., Z. (2006). Structural insights into histone demethylation by jmjd2 family members. *Cell* 125, 691–702. doi:10.1016/j.cell.2006.04.024

Cohen S, L. G. B. N. R. P. R. R., Graham M (2007). Protein composition of catalytically active human telomerase from immortal cells. *Science* 315, 1850–1853

Cui, Q., Elstner, M., Kaxiras, E., Frauenheim, T., and Karplus, M. (2001). A qm/mm implementation of the self-consistent charge density functional tight binding (scc-dftb) method. *Journal of Physical Chemistry B* 105, 569–585. doi:10.1021/jp0029109

DeChancie, H. K., J. (2007). The origins of femtomolar proteinligand binding: hydrogen-bond cooperativity and desolvation energetics in the biotin(strept)avidin binding site. *J. Am. Chem. Soc.* 129, 5419–5429

Du, M.-H., Kolchin, A., and Cheng, H.-P. (2004). Hydrolysis of a two-membered silica ring on the amorphous silica surface. *The Journal of Chemical Physics* 120, 1044–1054. doi:10.1063/1.1630026

Feynman, R. (1982). Simulating physics with computers. *Int. J. Theor. Phys.* 21, 467–488

Friesner, R. A. and Guallar, V. (2005). Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (qm/mm) methods for studying enzymatic catalysis. *Annual Review of Physical Chemistry* 56, 389–427. doi:10.1146/annurev.physchem.55.091602.094410

Gao, H. J. and Klein, P. (1998). Numerical simulation of crack growth in an isotropic solid with randomized internal cohesive bonds. *Journal of the Mechanics and Physics of Solids* 46, 187–218. doi:10.1016/s0022-5096(97)00047-1

Gao, J. L., Amara, P., Alhambra, C., and Field, M. J. (1998). A generalized hybrid orbital (gho) method for the treatment of boundary atoms in combined qm/mm calculations. *Journal of Physical Chemistry A* 102, 4714–4721. doi:10.1021/jp9809890

Gao, J. L. and Xia, X. F. (1992). A priori evaluation of aqueous polarization effects through monte-carlo qm-mm simulations. *Science* 258, 631–635. doi:10.1126/science.1411573

Georges, A., Kotliar, G., Krauth, W., and Rozenberg, M. J. (1996). Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Reviews of Modern Physics* 68, 13–125. doi:DOI10.1103/RevModPhys.68.13

Gordon, M. S., Fedorov, D. G., Pruitt, S. R., and Slipchenko, L. V. (2012). Fragmentation methods: A route to accurate calculations on large systems. *Chemical Reviews* 112, 632–672. doi:10.1021/cr200093j

Grimsley, H. R., Economou, S. E., Barnes, E., and Mayhall, N. J. (2019a). An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications* 10, 3007. doi:10.1038/s41467-019-10988-2

Grimsley, H. R., Economou, S. E., Barnes, E., and Mayhall, N. J. (2019b). An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* 10, 3007

Gull, E., Millis, A. J., Lichtenstein, A. I., Rubtsov, A. N., Troyer, M., and Werner, P. (2011). Continuous-time monte carlo methods for quantum impurity models. *Rev. Mod. Phys.* 83, 349–404. doi:10.1103/RevModPhys.83.349

Harvey, R. (2014). *Biochemistry, Lippincott's Illustrated Reviews* (Massachusetts: Wolters Kluwer Health)

Hedin, L. (1965). New method for calculating 1-particle greens function with application to electron-gas problem. *Physical Review* 139, A796–+. doi:10.1103/PhysRev.139.A796

Hoekstra, A., Chopard, B., and P., C. (2014). Multiscale modelling and simulation: a position paper. *Phil. Trans. R. Soc. A* 372, 20130377. doi:10.1098/rsta.2013.0377

526 Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., et al. (2017). Hardware-
527 efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* 549,
528 242–246
529 Karplus, M. and McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature*
530 *Structural Biology* 9, 646–652. doi:10.1038/nsb0902-646
531 Kitaev, A. (1995). Quantum measurements and the abelian stabilizer problem. ArXiv:quant-ph/9511026
532 Laio, A., VandeVondele, J., and Rothlisberger, U. (2002). A hamiltonian electrostatic coupling scheme for
533 hybrid car-parrinello molecular dynamics simulations. *Journal of Chemical Physics* 116, 6941–6947.
534 doi:10.1063/1.1462041
535 Leonov, I., Anisimov, V. I., and Vollhardt, D. (2015). Metal-insulator transition and lattice instability of
536 paramagnetic v_2o_3. *Phys. Rev. B* 91, 195115. doi:10.1103/PhysRevB.91.195115
537 Liu, B., Huang, Y., Jiang, H., Qu, S., and Hwang, K. C. (2004). The atomic-scale finite element method.
538 *Computer Methods in Applied Mechanics and Engineering* 193, 1849–1864. doi:10.1016/j.cma.2003.
539 12.037
540 McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C., and Yuan, X. (2020). Quantum computational
541 chemistry. *Rev. Mod. Phys.* 92, 015003 (59 pages). ArXiv 1808.10402
542 Mitchell, G. A. F. M. F. H. S. E., M. (2010). Structural basis for telomerase catalytic subunit tert binding
543 to rna template and telomeric dna. *Nat Struct Mol Biol* 17, 513–518
544 Nam, Y., Chen, J.-S., Pisenti, N. C., Wright, K., Delaney, C., Maslov, D., et al. (2020). Ground-state
545 energy estimation of the water molecule on a trapped-ion quantum computer. *npj Quantum Information*
546 6, 33
547 Ng, K. K. M. M. B. D. P. E. L. B. B. J. S. P. G. O. v. D. F. R. N. O. J. S. J. B. T. S. M. S. C. O. U., S.S.
548 (2007). Crystal structures of histone demethylase jmjd2a reveal basis for substrate specificity. *Nature*
549 448, 87–91. doi:10.1038/nature05971
550 Pedersen, M. and Helin, K. (2010). Histone demethylases in development and disease. *Trends Cell Biol.*
551 20, 662–671. doi:10.1016/j.tcb.2010.08.011
552 Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., et al. (2014). A variational
553 eigenvalue solver on a photonic quantum processor. *Nat. Commun.* 5, 4213
554 Petersilka, M., Gossmann, U. J., and Gross, E. K. U. (1996). Excitation energies from time-dependent
555 density-functional theory. *Phys. Rev. Lett.* 76, 1212–1215. doi:10.1103/PhysRevLett.76.1212
556 Reiher, M., Wiebe, N., Svore, K. M., Wecker, D., and Troyer, M. (2017). Elucidating reaction mechanisms
557 on quantum computers. *Proceedings of the National Academy of Sciences* 114, 7555–7560. doi:10.
558 1073/pnas.1619152114
559 [Dataset] Roe, D., Okur, A., Simmerling, C., and Walker, R. (2008). Tutorial A7: Replica Exchenge
560 Rountree, C. L., Kalia, R. K., Lidorikis, E., Nakano, A., Van Brutzel, L., and Vashishta, P. (2002).
561 Atomistic aspects of crack propagation in brittle materials: Multimillion atom molecular dynamics
562 simulations. *Annual Review of Materials Research* 32, 377–400. doi:10.1146/annurev.matsci.32.111201.
563 142017
564 Rudd, R. E. and Broughton, J. Q. (2000). Concurrent coupling of length scales in solid state systems.
565 *Physica Status Solidi B-Basic Research* 217, 251–291. doi:10.1002/(sici)1521-3951(200001)217:
566 1(251::Aid-pssb251)3.0.Co;2-a
567 Sadava, H. D. H. C. . B. M., D. (2011). *Life: The science of biology* (Sunderland, MA: Sinauer Associates
568 Inc.)
569 Salomon-Ferrer, R., Case, D. A., and Walker, R. C. (2013). An overview of the Amber biomolecular
570 simulation package. *WIREs Comput. Mol. Sci.* 3, 198–210. doi:10.1002/wcms.1121
571 Seabra, G. d. M., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007). Implementation of
572 the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package.
573 *J. Phys. Chem A* 111, 5655–5664. doi:10.1021/jp070071l
574 Talebi, H., Silani, M., Bordas, S. P. A., Kerfriden, P., and Rabczuk, T. (2014). A computational library
575 for multiscale modeling of material failure. *Computational Mechanics* 53, 1047–1071. doi:10.1007/
576 s00466-013-0948-2
577 Tanaka, S., Mochizuki, Y., Komeiji, Y., Okiyamac, Y., and Fukuzawace, K. (2014). Electron-correlated
578 fragment-molecular-orbital calculations for biomolecular and nano systems. *Phys. Chem. Chem. Phys.*
579 16, 10310–10344. doi:10.1039/c4cp00316k
580 Tran, L. N., Iskakov, S., and Zgid, D. (2018). Spin-unrestricted self-energy embedding theory. *The Journal*
581 *of Physical Chemistry Letters* 9, 4444–4450. doi:10.1021/acs.jpclett.8b01754. PMID: 30024163

582 Vreven, T., Morokuma, K., Farkas, O., Schlegel, H. B., and Frisch, M. J. (2003). Geometry optimization
583 with qm/mm, oniom, and other combined methods. i. microiterations and constraints. *Journal of*
584 *Computational Chemistry* 24, 760–769. doi:10.1002/jcc.10156
585 Walsh, T. R., Wilson, M., and Sutton, A. P. (2000). Hydrolysis of the amorphous silica surface. ii.
586 calculation of activation barriers and mechanisms. *The Journal of Chemical Physics* 113, 9191–9201.
587 doi:10.1063/1.1320057
588 Witzany, G. (2008). The viral origins of telomeres, telomerases and their important role in eukaryogenesis
589 and genome maintenance. *Biosemiotics* 1,191–206
590 Zahariev, F. and Gordon, M. S. (2012). Development of a combined quantum Monte Carlo-effective
591 fragment molecular orbital method. *Molecular Physics* 117, 1532–1540. doi:10.1080/00268976.2019.
592 1574363
593 Zhang, L., Staar, P., Kozhevnikov, A., Wang, Y. P., Trinastic, J., Schulthess, T., et al. (2019a). Dft plus
594 dmft calculations of the complex band and tunneling behavior for the transition metal monoxides mno,
595 feo, coo, and nio. *Physical Review B* 100, 035104. doi:10.1103/PhysRevB.100.035104
596 Zhang, L., Staar, P., Kozhevnikov, A., Wang, Y.-P., Trinastic, J., Schulthess, T., et al. (2019b). DFT +
597 DMFT calculations of the complex band and tunneling behavior for the transition metal monoxides
598 mno, feo, coo, and nio. *Phys. Rev. B* 100, 035104. doi:10.1103/PhysRevB.100.035104