Quantum Simulation of Preferred Tautomeric State Prediction

Yu Shee,[1, 2, *] Tzu-Lan Yeh,[3, *] Jen-Yueh Hsiao,[4] Ann Yang,[3] Yen-Chu Lin,[3, 5, †] and Min-Hsiu Hsieh[4, ‡]

[1] *Bleximo Corp., Berkeley, California 94720, USA*
[2] *Department of Chemistry, Yale University, New Haven, Connecticut 06511, USA*
[3] *Insilico Medicine Taiwan Ltd., Taipei City 110208, Taiwan*
[4] *Hon Hai (Foxconn) Research Institute, Taipei City 114, Taiwan*
[5] *Department of Pharmacy, National Yang Ming Chiao Tung University, Taipei City 112304, Taiwan*
(Dated: October 7, 2022)

Prediction of tautomers plays an essential role in computer-aided drug discovery. However, it remains a challenging task nowadays to accurately predict the canonical tautomeric form of a given drug-like molecule. Lack of extensive tautomer databases, most likely due to the difficulty in experimental studies, hampers the development of effective empirical methods for tautomer predictions. A more accurate estimation of the stable tautomeric form can be achieved by quantum chemistry calculations. Yet, the computational cost required prevents quantum chemistry calculation as a standard tool for tautomer prediction in computer-aided drug discovery. In this paper we propose a hybrid quantum chemistry-quantum computation workflow to efficiently predict the dominant tautomeric form. Specifically, we select active-space molecular orbitals based on quantum chemistry methods. Then we utilize efficient encoding methods to map the Hamiltonian onto quantum devices to reduce the qubit resources and circuit depth. Finally, variational quantum eigensolver (VQE) algorithms are employed for ground state estimation where hardware-efficient ansatz circuits are used. To demonstrate the applicability of our methodology, we perform experiments on two tautomeric systems: acetone and Edaravone, each having 52 and 150 spin-orbitals in the STO-3G basis set, respectively. Our numerical results show that their tautomeric state prediction agrees with the CCSD benchmarks. Moreover, the required quantum resources are efficient: in the example of Edaravone, we could achieve chemical accuracy with only eight qubits and 80 two-qubit gates.

I. INTRODUCTION

Tautomers are constitutional isomers that spontaneously convert to one another in dynamic equilibrium. The process of this interconversion is called tautomerization. Typical tautomerization involves the movement of a proton from one position to another and rearrangement of a double bond within the molecule. Other types of tautomerisms include annular, ring-chain, and valence tautomerisms [1–3]. One well-known example of tautomerization, and quite often involved in the field of drug development, is the keto-enol tautomerism, in which the carbonyl double bond (keto form) is interconverted to an alkene double bond (enol form). This is accompanied by the shift of the alpha proton in the keto form to the hydroxyl group in the enol form.

Tautomerization plays an important role in biological systems. Non-Watson-Crick base pairing can occur due to tautomerization of nucleic acid base pairs. Such nucleic acid mismatches induced by tautomerization result in spontaneous mutagenesis and hence genetic instability [4, 5]. In the case of drug molecules, the movement of proton within the molecules will lead to the conversion

---

* These authors contributed equally to this work.
† Corresponding e-mail: jimmy.lin@insilicomedicine.com
‡ Corresponding e-mail: min-hsiu.hsieh@foxocnn.com

from a hydrogen bond donor to a hydrogen bond acceptor or vice versa, which is essential for the analysis of structure-activity relationship (SAR). Moreover, it has been estimated that more than a quarter of marketed drugs can exhibit tautomerism [6]. Prediction of tautomeric states of compounds of interest is therefore an important subject in the field of computer aided drug design.

State of the art algorithms for tautomeric prediction usually involve enumeration of possible tautomers, followed by prediction of the dominant form or estimation of the tautomer populations [7]. The latter is usually achieved using empirical rules in combination with scoring functions or relies on the estimation of pKa, etc. Such algorithms usually aim to provide a list of possible tautomeric forms with ratios of the corresponding species, since it is important to estimate the population of the tautomers due to the small differences in their free energies that could be easily compensated by the interaction with proteins. Lack of extensive data on free energies between tautomers and that in different solvent environments render a great challenge for accurate prediction by empirical methods. It is thus our hope that quantum chemical approaches could provide better accuracy and fill the gap.

Computational quantum chemistry algorithms are heavily used in drug design and discovery to simulate chemical reactions and properties. These quantum chemistry computational strategies can accelerate the drug discovery processes and reduce the costs for wet lab experiments. While quantum chemistry approaches have brought accuracy to the calculations, their computational cost makes it not suitable as a daily used tool in computer-aided drug discovery (CADD). Instead, empirical or semi-empirical methods are used as standard tools. Emerging use of density functional calculations brings balance between accuracy and computation cost in the field. As such, density functional theory (DFT) methods are most widely used for the calculation of electron density of molecular systems. However, their accuracy depends on the electronic structures of the molecular systems; hence, DFT is not universally applicable [8]. Instead, wave function based techniques, such as the configuration interaction (CI) and coupled cluster (CC) methods, are used. The runtime of the full CI (FCI) method, which is equivalent to diagonalizing the Hamiltonian matrix, scales exponentially with respect to system sizes. The coupled cluster singles and doubles (CCSD) method, on the other hand, provides reasonable trade-offs between accuracy and runtime. Even though the aforementioned classical algorithms are well-designed, they still have unreasonable time and space requirements for medium and large molecules [9]. This poses severe limitation on the use of quantum chemistry methods in drug discovery, e.g., prediction of dominant product in tautomerism, understanding the dynamics of protein folding, and ligand binding free energy calculations.

Quantum simulation is considered one of the most promising applications of quantum technology since it has the potential of overcoming the exponential barrier of solving electronic structure problems [10]. Jordon-Wigner (JW) transformation [11] was found to give mappings from fermionic creation and annihilation operators to qubit operators [12]. This result paved the road to generically simulate physical systems on well-controlled quantum computers. Moreover, the quantum phase estimation (QPE) algorithm [13–22] provides accurate spectral calculations for molecular Hamiltonian with the potential of exponential speedup. However, the QPE method is not applicable on current quantum devices because it requires long coherence time and high gate fidelity. Instead, several variational methods [17, 23–33] suitable for noisy intermediate-scale quantum (NISQ) devices were proposed. One of the most practical approaches is the variational quantum eigensolver (VQE) algorithm [23, 24]. VQE is a hybrid quantum-classical method that approximates ground states of molecular Hamiltonian by variationally tuning the ansatz parameters and is expected to give better accuracy than CCSD results with polynomial costs.

Quantum simulation of small molecules has since then been widely implemented. Researchers have worked towards molecules (BeH_2, H_2O, and H_{12}) with improved algorithms and hardware design [25, 34, 35]. Others also provide strategies to simulate slightly larger systems, such as

CO_2, C_2H_4, C_{18}, and the nitrogenase iron-sulfur molecular clusters [36–38], using symmetries, fragmentation of molecules, spin-model simplifications, or novel qubit encoding methods [39–44]. However, these quantum simulation results are still quite limited in the problem size and often lack real world applications.

In this work, we aim to design a general methodology for a pharmaceutical application; namely, the prediction of preferred tautomeric states. We expect that our scheme could work for both current and future quantum devices. Specifically, we presented a hybrid quantum chemistry-quantum computation approach to predict the dominant form of tautomers, where quantum chemistry methods are used to construct the system of interest and reduce the size of the Hamiltonian to meet current quantum hardware requirement, and the reduced Hamiltonian was mapped to qubits using qubit-efficient encoding (QEE) [44] and subsequently simulated with variational quantum eigensolver (VQE). We focus on the problem of tautomer ratio estimation and prediction of the preferred tautomeric state by considering the energetics (relative free energy) of the states using quantum chemistry approaches to make the system applicable on current quantum cumputing schemes. This approach was practiced on two illustrative examples, acetone and Edaravone. Our simulation results showed reasonable agreement with the calculations using CCSD.
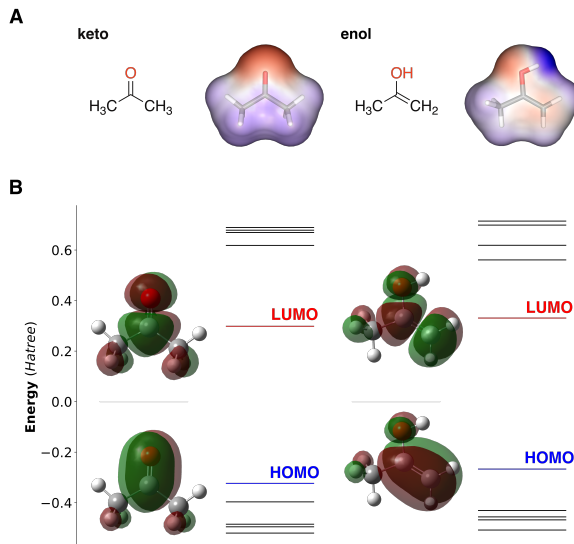
II. RESULTS

FIG. 1: (A) Tautomeric isomers of acetone. The geometries were optimized at the B3LYP/6-311++G(d,p) level and the electrostatic potential surface generated by ESP_DNN[45]. (B) Energy levels of molecular orbitals around HOMO/LUMO as calculated using MP2/STO-3g. The full list molecular orbital eignevalues can be found in Appendix B.

The most simple and intuitive way to estimate the relative stability of tautomers is perhaps by considering the difference in bond energy, since tautomerization usually involves rearrangement of double bonds. Take the well-known keto-enol tautomerism for example. The bond dissociation energy difference between C=O, C-C, C-H (keto form) and O-H, C-O, C=C (enol form) is 66 kJ/mol (15.8 kcal/mol), favoring the keto form. However, as the complexity of the molecules increases, the relative stability could not be attributed so easily. Factors like electrostatic effect, steric hindrance from the other parts of the molecule and intra-molecular hydrogen bonds need careful consideration as well. Commercial or publicly available CADD tools deal with the tautomerism problem using empirical or rule-based chemoinformatics methodologies, e.g., scoring tautomers based on the prediction of microstate and microstate pKa values. There is, however, still room for improvement with such an approach due to the lack of extensive databases and over-parameterization of the prediction models.

In principle, two important factors decide the tautomeric equilibrium. The relative potential energy difference, which determines the equilibrium direction, accounts for the stability between the isomers. On the other hand, the rate of isomer interconversion depends on the activation energy. As tautomerism involves bond breaking and bond formation, the best way to handle this problem should be quantum mechanics-based approaches with the system described by molecular orbitals. However, such first-principle calculations are impractical even with the computation resources nowadays. With the advent of quantum computing technologies, theoretical quantum chemical approaches could now play an important and 'applicable' role for the tautomer prediction for computer aided drug discovery.

In the following, we first present the quantum chemistry overview of the two tautomeric systems: (1) acetone and its enol form, and (2) Edaravone's keto, enol, and amine forms in the STO-3G basis set. Next, we introduce current challenges of simulating medium-to-large molecules and our workflow that could contribute to resolving the problems. Besides, we provide the background of the qubit-efficient encoding methods for quantum simulation. Lastly, we show the numerical results for the implementation of our workflow with quantum simulations of the two tautomeric systems. To the best of our knowledge, these are the largest molecules in terms of the number of spin-orbitals (without the fragmentation of molecules using quantum embedding theory or the application of spin models from spectroscopic studies) for numerical simulation using a quantum algorithm so far.

Overview of Systems

Acetone. We start with a relatively simple example of tautomerization. Acetone exhibits keto-enol tautomerization, in which the acetone and its enol form, propen-2-ol, interconvert to each other (Fig. 1(A)). The relative stability of these two isomers can be easily assessed by comparing their bond energy, and the keto from is more stable than the enol form by about 16 kcal/mol. Fig. 1(A) shows the optimized geometries of the keto and enol forms of the acetone at the level of 6-311++G(d,p) using B3LYP in the Gaussian 16 program. To enable quantum computing of the system at a reasonable computational cost, we reconstituted the molecules with the minimal STO-3G at the second-order Møller–Plesset perturbation theory (MP2) level with 52 spin-orbitals. The corresponding energy diagrams of the molecular orbitals near HOMO are shown in Fig 1(B) for the keto and enol form, respectively. The keto form is more stable than the enol form in most cases because the C=O double bond is stronger than the C=C bond. Polarization of the C=O double bond, as depicted in electrostatic potential surface of the acetone, gives it a relatively higher bond energy. The rearrangement of the double bond between the keto and enol form is apparent from the shift of the π bonding patterns of the HOMOs for the corresponding tautomers.
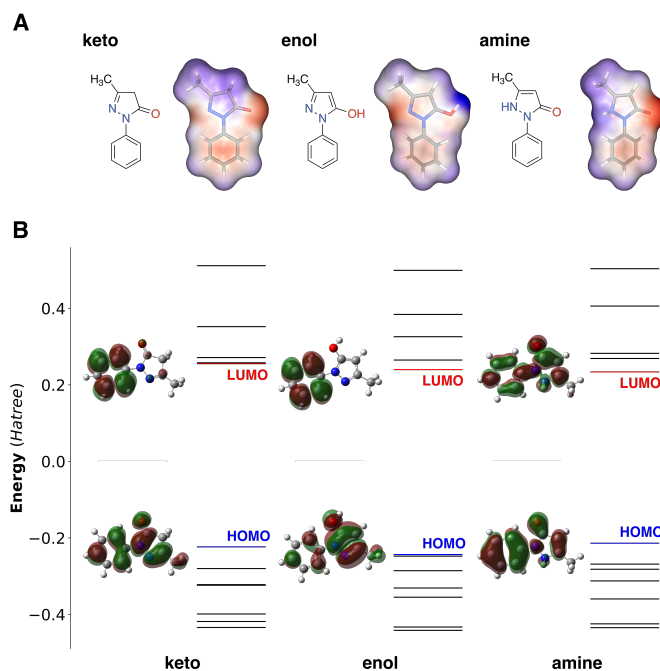
FIG. 2: (A)Tautomeric isomers of Edaravone. The geometries were optimized at the B3LYP/6-311++G(d,p) level and the electrostatic potential surface generated by ESP_DNN [45]. (B) Energy levels of molecular orbitals around HOMO/LUMO as calculated using MP2/STO-3g. The full list molecular orbital eignevalues can be found in Appendix B.

Edaravone. Edaravone is an FDA-approved drug for the treatment of amyotrophic lateral sclerosis (ALS). The mechanism of the action of Edaravone involves its radical scavenging activity. The idea originated from a research program in Mitsubishi Yuka Pharmaceutical Corporation, in which the scientists try to utilize the radical-scavenging activity of phenol and avoid related toxicity. The researchers believed that an aromatic heterocyclic system with the potential of keto-enol tautomerization could exist in the form with a hydroxyl group [46]. Edaravone exists in three forms of tautomeric isomers – keto, enol, and amine (Fig. 2A), with varying physicochemical properties. Moreover, the three forms of Edaravone exhibit different antioxidant activity. Therefore, it is important to know which form among the three dominates to better understand the pharmacological effect of Edaravone. Estimation of relative stability among the three tautomers for Edaravone is not as trivial as for acetone, since the molecular structure of Edaravone is much more complicated. For example, the aromaticity of the N-substituted pyrazolone core in the three isomers varies. Moreover the surface charge distribution of the three tautomers differs quite a lot and the twisted angles between the pyrazolone and the N-substituted benzene are distinct among the three according to the optimized geometries of the forms (at the level of B3LYP/6-311++G(d,p) theory) as depicted in Fig 2(A). Reconstitution of the electronic configuration using the minimal STO-3g basis at MP2 level yields 150 molecular orbitals. Energy diagrams of the molecular orbitals near HOMO or the

HOMOs and LUMOs of the three tautomeric species do not provide clear pictures for the relative stabilities of the three species as in the case of acetone (Fig. 2(B)).

Notice that it is already considered computationally expensive to construct the electronic structures of Edavarone even with classical quantum chemistry approaches. However, in the field of medicinal chemistry, there are more compounds with larger sizes. It is therefore not an efficient way to predict the thermodynamic stability of different tautomer states by comparing the energies of the isomers using quantum chemistry directly.

In order to leverage the power of quantum computation, we proceed as follows. Firstly, the second-quantized electronic Hamiltonian can be written as

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \tag{1}$$

where h_{pq} and h_{pqrs} are the overlap and exchange integrals. The indices of the spin-orbitals are represented by p, q, r, and s in the summation of Eq. (1). For example, the indices run from 0 to 51 for acetone and from 0 to 149 for Edaravone because the molecules in the STO-3G basis set have 52 and 150 spin-orbitals respectively. To simulate electronic structure problems on quantum computers, the creation and annihilation operators and the electronic states have to be mapped to qubits.

Common qubit encoding methods are Jordan-Wigner, parity, and Bravyi-Kitaev encoding where the qubit requirements are $\mathcal{O}(N)$ where N is the number of spin-orbitals. For example, Jordan-Wigner (JW) encoding uses N qubits to store the N spin-orbital occupation number. In this encoding scheme, the $|0\rangle$ qubit state represents that the spin-orbital is not occupied while the $|1\rangle$ qubit state represents that the spin-orbital is occupied. The mappings of the creation and annihilation operators for JW are

$$a_p^\dagger = \frac{1}{2}(X_p - iY_p) \otimes Z_{p-1} \otimes \cdots \otimes Z_0, \tag{2}$$

$$a_p = \frac{1}{2}(X_p + iY_p) \otimes Z_{p-1} \otimes \cdots \otimes Z_0, \tag{3}$$

where the Z Pauli strings address the fact that electronic wavefunctions are anti-symmetric. For acetone and Edaravone in the STO-3G basis set, JW encoding will map the electronic structure problems to 52 qubits and 150 qubits. No meaningful data would be collected for VQE using that many qubits on NISQ devices. It is, therefore, important to reduce the problem size.

Workflow

Here we propose a hybrid computational scheme for the prediction of the preferred tautomeric state, comprising classical quantum chemistry (using classical computers) and quantum computation. The workflow involves (1) geometry optimization of the isomers using density functional theory, (2) reconstruction of the molecular orbital space in a minimal basis set, (3) reduction of the MO space, (4) single point energy calculations of active MO sets, (5) selection of the most representative active MO set, and finally (6) energy calculation in Quantum computing scheme using QEE, as illustrated in Fig. 3. We use the CCSD energies of the full systems as the benchmark of our workflow performance.

Specifically, to reduce the problem size, we use quantum chemistry methods to select active molecular orbitals (MOs) and map the Hamiltonian onto fewer qubits using a qubit-efficient encoding method. The geometries of the tautomeric isomers are prepared and initially optimized
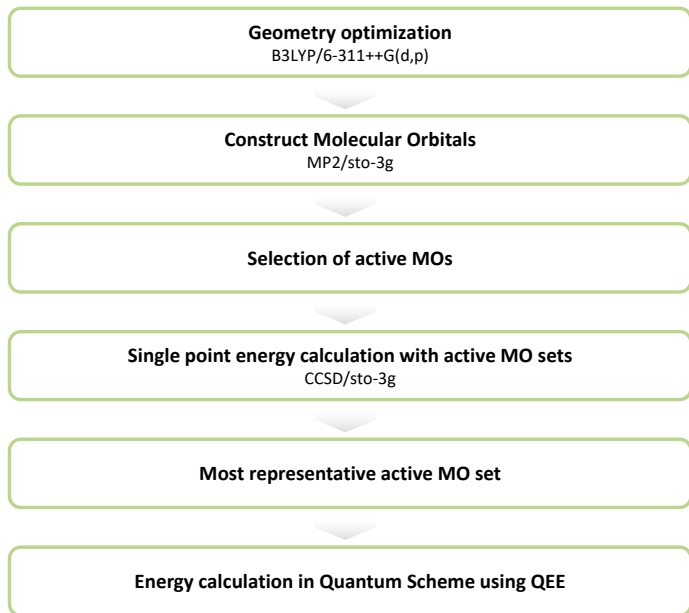
FIG. 3: Proposed workflow of preferred tautomeric state prediction

using Avogadro [47] with the MMFF94 force field [48–52] and further optimized at the level of 6-311++G(d,p) using B3LYP [53, 54] in the Gaussian 16 program. Due to limited resources for quantum computing, the molecules are reconstituted with the minimal STO-3g at the MP2 level [55], followed by the calculation of occupancies of natural orbitals (NOs) using Gaussian 16. The STO-3g basis set has been widely used in several pioneering works of computation chemistry with quantum computing scheme [25, 34, 35]. Although the computational cost has been greatly reduced with the minimal basis set, the number of required qubits still exceeds the affordable amount of current QC simulators or real quantum devices. The reduction of the MO space is then achieved by considering the occupation numbers and quantum hardware limitations. The selection of active space was achieved mostly by excluding the inner core and the outer virtual orbitals. In practice, the natural orbitals with occupancy closest to 0 (completely virtual) or 2 (fully occupied) were frozen or removed. Besides, quantum hardware constraints and the efficiencies of the qubit-efficient encoding method are also considered in the selection process. The reduced second-quantized Hamiltonian is then encoded onto qubits.

The natural orbitals for Acetone and Edaravone tatutomers calculated at the MP2/STO-3G level can be found in Appendix B. Selected combinations of active MOs are listed in Appendix C for acetone and Edaravone, respectively. After the active space selection from quantum chemistry methods, we have 12 spin-orbitals (6 molecular orbitals) for acetone, propen-2-ol, and the three tautomers of Edaravone. With Jordan-Wigner (JW) encoding, the qubit count to simulate the quantum systems is 12 qubits which is still a difficult task for noisy intermediate-scale quantum devices without error mitigation. Therefore, we use qubit-efficient encoding (QEE) to encode the systems onto 8 qubits shown in Table I because it provides fewer qubit counts and renders the systems to be suitable for the use of hardware-efficient ansatzes. We then use a hardware-efficient ansatz to simulate the systems with variational quantum eigensolver (VQE).

Molecule	Active Space Orbitals	(Electrons, MO)	JW Qubit Count	QEE Qubit Count
acetone	14-19	(4e, 6o)	12	8
propen-2-ol	14-19	(4e, 6o)	12	8
Edaravone (keto)	44-49	(4e, 6o)	12	8
Edaravone (enol)	44-49	(4e, 6o)	12	8
Edaravone (amine)	44-49	(4e, 6o)	12	8

TABLE I: Comparison of qubit counts for Jordan-Wigner encoding and qubit-efficient encoding for acetone, propen-2-ol, and the three tautomers of Edaravone (in the STO-3G basis sets and equilibrium bond distances) with some molecular orbitals being frozen/removed. Note that the molecular orbital indices are ordered from the lowest to the highest energies and the indices start from 0.

Even though the qubit requirement is reduced by classical active space selection methods, it is still hard to be implemented on a quantum device as the circuit depth would not permit. For JW encoding, some unphysical (e.g. not particle conserving or violating other symmetries) electronic configurations are also encoded. This necessitates the usage of chemical-inspired ansatzes (e.g. the unitary couple cluster (UCC) ansatz) that often have larger circuit depth so that the trial wavefunction only represents the states in the chemical subspace. It is still possible to use a hardware-efficient ansatz for JW encoded problems, but the number of entangling layers has to be large. With more entangling layers, the ansatz circuit would have better expressibility [56] and entanglement to include the solution space [57], but it may suffer from the vanishing gradient and the barren plateau problem [58–60]. Thus, we resort to a qubit-efficient encoding (QEE) method proposed in [44] which provides an logarithmic saving of qubit counts (8 qubits for QEE). Additionally, QEE only encodes the physical (significant) electronic configurations such that low depth hardware-efficient ansatzes are suitable for QEE encoded problems.

Qubit Efficient Encoding. As we want to work on a larger electronic system, an encoding scheme using $\mathcal{O}(N)$ qubits would not be practical on current and near-future devices. Therefore, it would be suitable to resort to encoding methods using fewer qubits. Here, we choose Qubit Efficient Encoding (QEE) from [44] where the qubit scales logarithmically with respect to N. This is done by only mapping particle conserving or other eletronic configurations with certain symmetries to the qubit basis states which exploits the exponential growth of qubit Hilbert space.

Since the encoding scheme works in a subspace of the space spanned by the occupation basis, we will not be able to map a single creation or annihilation operator as the operators change the number of electrons in the system. Nevertheless, the second-quantized Hamiltonian can be written as a linear combination of excitation operators, which conserve the number of particles, $E_{pq} \equiv a_p^\dagger a_q$ yielding

$$H = \sum_{pq} h_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} h_{pqrs} (\delta_{qr} E_{ps} - E_{pr} E_{qs}) \tag{4}$$

by using the anti-commutation relations of fermionic operators. Any excitation operator E_{pq} can be written as $E_{pq} = \sum_{k,k'=0}^{|\mathcal{F}_m|-1} c_{k'k}^{pq} |\mathbf{f}_{k'}\rangle_\mathrm{f} \langle \mathbf{f}_k|_\mathrm{f}$, where $\mathcal{F}_m$ is the set of all particle conserving (or with other symmetries) electronic configurations $|\mathbf{f}_k\rangle_\mathrm{f}$ and $c_{k'k}^{pq} = \langle \mathbf{f}_{k'}|_\mathrm{f} E_{pq} |\mathbf{f}_k\rangle_\mathrm{f}$. $c_{k'k}^{pq}$ is zero if the transition from $|\mathbf{f}_k\rangle_\mathrm{f}$ to $|\mathbf{f}_{k'}\rangle_\mathrm{f}$ via excitation operator E_{pq} is impossible. The coefficient $c_{k'k}^{pq}$ can be ± 1 according to the antisymmetric property of fermions where $c_{k'k}^{pq} = \prod_{i=\min(p,q)+1}^{\max(p,q)-1} (-1)^{f_i}$. With a transformation $\mathcal{E}$ that maps the selected electronic configurations one-to-one to qubit basis states

in an arbitrary order, the corresponding qubit operator of E_{pq} can be written as

$$\tilde{E}_{pq} = \mathcal{E} \circ E_{pq} \circ \mathcal{E}^{-1} = \sum_{k,k'=0}^{|\mathcal{F}_m|-1} c_{k'k}^{pq} |\mathbf{q}_{k'}\rangle_{\mathrm{q}} \langle \mathbf{q}_k|_{\mathrm{q}}, \tag{5}$$

where $|\mathbf{q}_k\rangle_{\mathrm{q}} = \mathcal{E} |\mathbf{f}_k\rangle_{\mathrm{f}}$ is the encoded qubit state of $|\mathbf{f}_k\rangle_{\mathrm{f}}$ under the map $\mathcal{E}$. The transition between the two basis states $|\mathbf{q}\rangle_{\mathrm{q}} = |q_{Q-1}, ..., q_0\rangle_{\mathrm{q}}$ and $|\mathbf{q}'\rangle_{\mathrm{q}} = |q'_{Q-1}, ..., q'_0\rangle_{\mathrm{q}}$ can be factorized as $|\mathbf{q}'\rangle_{\mathrm{q}} \langle \mathbf{q}|_{\mathrm{q}} = \bigotimes_{k=0}^{Q-1} |q'_k\rangle_{\mathrm{q}} \langle q_k|_{\mathrm{q}}$. Therefore, the encoded excitation operator can be expressed as

$$\tilde{E}_{pq} = \sum_{k,k'=0}^{|\mathcal{F}_m|-1} c_{k'k}^{pq} |\mathbf{q}_{k'}\rangle_{\mathrm{q}} \langle \mathbf{q}_k|_{\mathrm{q}} = \sum_{k,k'=0}^{|\mathcal{F}_m|-1} \bigotimes_{w=0}^{Q-1} c_{k'k}^{pq} T_{k'k,w}, \tag{6}$$

where $T_{k'k,w}$ could be one of the following operators

$$Q^+ = |1\rangle_{\mathrm{q}} \langle 0|_{\mathrm{q}} = \frac{1}{2}(X - iY), \tag{7}$$

$$Q^- = |0\rangle_{\mathrm{q}} \langle 1|_{\mathrm{q}} = \frac{1}{2}(X + iY), \tag{8}$$

$$N^{(0)} = |0\rangle_{\mathrm{q}} \langle 0|_{\mathrm{q}} = \frac{1}{2}(I + Z), \tag{9}$$

$$N^{(1)} = |1\rangle_{\mathrm{q}} \langle 1|_{\mathrm{q}} = \frac{1}{2}(I - Z) \tag{10}$$

according to the qubit basis state transitions.

As each operator $T_{k'k,w}$ is a sum of two Pauli operators, this allows us to express the qubit Hamiltonian

$$H_{\mathrm{q}} = \sum_{pq} h_{pq} \tilde{E}_{pq} + \frac{1}{2} \sum_{pqrs} h_{pqrs} (\delta_{qr} \tilde{E}_{ps} - \tilde{E}_{pr} \tilde{E}_{qs}) \tag{11}$$

as a sum of Pauli operator strings. With the Hamiltonian in Pauli strings, the energy expectation value can be obtained using a quantum computer. In this work, we select electronic configurations that are particle-conserving and singlet so we reduce the qubit requirement from 12 qubits to 8 qubits.

Variational quantum circuits and parameter optimization. After obtaining the qubit Hamiltonian from QEE, we employ two different hardware-efficient ansatz, as illustrated in Fig. 4, for the VQE algorithm. The first hardware-efficient ansatz consists of four two-qubit entangling building blocks , and the layers are arranged in a staggered form where Fig. 4a gives an example of two hardware-efficient layers. The first ansatz used in this work consists of at most 20 layers and eight parameterized R_y rotations at each layer in the end. Therefore, the largest circuit in this form has 80 CNOT gates and 168 parameters. While for the second/alternative ansatz, each layer consists of seven CNOT gates and eight parameterized R_y rotations shown in Fig. 4b. The largest circuit in this form consists of 10 layers and has 70 CNOT gates and 80 parameters. We have also run this alternative ansatz circuit on a noisy simulator where the circuit consists of four layers and has 28 CNOT gates and 32 parameters.

For the initial state, we have compared two initialization strategies. First, we have used the Hartree-Fock (HF) state as the initial state where all the parameters are set to zero initially [44].

We have also compared HF initialization with the Gaussian initialization strategy proposed in [61]. Zhang et al. [61] proposed this Gaussian initialization strategy to escape from barren plateau where the initial parameters are sampled from a Gaussian distribution. The Gaussian distribution used in this work has a mean of 0 and a variance of 0.3. The classical optimization of the variational parameters was done using the Sequential Least Squares Programming (SLSQP) method.

(a) Two hardware-efficient layers of the 8-qubit ansatz circuit. Note that there are single-qubit parametrized R_y rotations on all the qubits at the end of this ansatz circuit.

(b) One hardware-efficient layer of the alternative 8-qubit ansatz circuit. Note that there is no single-qubit rotation at the end of this ansatz circuit.
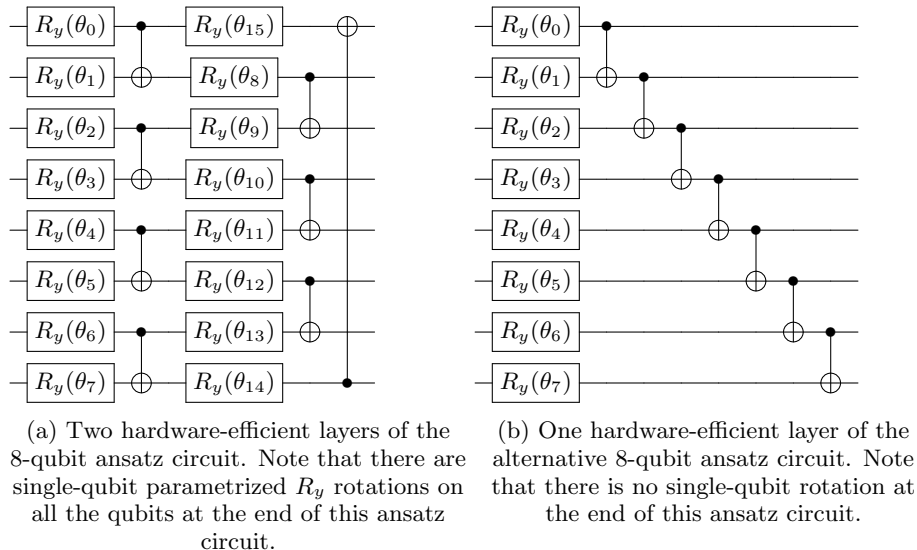
FIG. 4: The hardware-efficient ansatz circuit layers used in this work.

Numerical Results

We investigate the performance of our workflow with several numerical experiments. First, the CCSD energies for the full systems of the tautomers in the STO-3G basis set are calculated and are considered our benchmark for energy estimation performance. Second, energy calculations of VQE with hardware-efficient ansatz on a noiseless quantum simulator are performed for both the acetone and the Edaravone tautomers. Third, VQE energies on a noiseless quantum simulator for the acetone system using an alternative ansatz circuit are obtained. Lastly, a noisy simulation (with the noise model detailed in Appendix A) of the acetone system has been done using four layers of the alternative ansatz circuit.

The CCSD energies of the full systems and the active-space systems in the STO-3G basis set are computed. For the acetone system (Table II), the full system CCSD energy of the enol form relative to the keto form is 24.070 kcal/mol, and the active-space system CCSD energy of the enol form relative to the keto form is 23.766 kcal/mol. For the Edaravone system (Table III), the full system CCSD energy of the enol form relative to the keto form is 13.726 kcal/mol and the CCSD energy of the amine form relative to the keto form is 25.947 kcal/mol. While for active-space system of Edaravone, the CCSD energy of the enol form relative to the keto form is 9.197 kcal/mol and the CCSD energy of the amine form relative to the keto form is 25.398 kcal/mol.

For the VQE results of acetone shown in Fig. 5, we use two different parameter initialization

	acetone	Propen-2-ol
Full system CCSD energies relative to acetone	0	24.070
Active set CCSD energies relative to acetone	0	23.766

TABLE II: Relative ground state energy of acetone and propen-2-ol in kcal/mol. Relative energies for both the full MO and the active set (HOMO-1, HOMO, LUMO, LUMO+1, LUMO+2, and LUMO+3) were shown. Various combinations of orbital reduction are provided in Supplementary Information Table X.

	Keto	Enol	Amine
Full system CCSD energies relative to keto form	0.000	13.726	25.947
Active set CCSD energies relative to keto form	0.000	9.197	25.398

TABLE III: Relative ground state energy of the keto, enol, and amine form of Edavarone in kcal/mol. Relative energies for both the full MO and the active set (HOMO-1, HOMO, LUMO, LUMO+1, LUMO+2, and LUMO+3) were shown. Various combinations of orbital reduction are provided in Supplementary Information Table XI.

strategies (Gaussian and HF) and investigate how different number of ansatz circuit layers affect VQE calculations. For acetone in Fig. 5a with 20 circuit layers using Gaussian initialization, the keto form has an error of 0.135 kcal/mol and the enol form has an error of 0.258 kcal/mol which are well below chemical accuracy (1 kcal/mol). It can be seen that more entangling ansatz layers provides lower VQE error (see Fig. 5a) and better estimations of the relative energy gaps between the keto and enol forms of acetone (see Fig. 5b). Besides, VQE calculations of active-space systems have the same results with CCSD calculations of the full systems for the predictions of stability between the tautomers (see the dotted lines in Fig. 5b). However, for the case where initial states are HF states, the VQE results are often found to be trapped in local minimum or barren plateau. This can be clearly observed from the keto form of acetone in Fig. 5c where the final states are almost the same as the initial HF state. While for the case where Gaussian initialized parameters are used, the VQE shows faster convergence to chemical accuracy (require lower circuit depth) where only 12 circuit layers are needed to reach chemical accuracy. Also, the relative energy gaps between the keto and enol forms of acetone are not accurate when using HF intialization (see Fig. 5d).

Similar VQE results can be found in Fig. 6 for the Edavarone system. For Edaravone in Fig. 6a with 20 circuit layers, the keto form has an error of 0.730 kcal/mol, the enol form has an error of 1.006 kcal/mol, the amine form has an error of 1.168 kcal/mol which are fairly close to chemical accuracy. Fig. 6a also shows that more entangling ansatz layers often provides lower VQE error and better estimations of the relative energy (Fig. 6b). For the case of HF initialization for Edavarone, the VQE errors and calculations of relative energy gaps also reach similar accuracy in Fig. 6c and Fig. 6d. However, it can still be observed that Gaussian initialization has a better convergence with respect to the number of circuit layers.

Additionally, we tested the performance of an alternative ansatz (see Fig. 4b for a single layer) on the acetone system. The results in Fig. 7 show that the energy gap estimation has reached chemical accuracy with just 4 entangling layers (28 two-qubit gates). The circuit depth and the two-qubit gate count are reasonable for near-term quantum devices. Thus, we have run the circuit with the parameters being set at the optimal point on a noisy simulator (the noise model is based on the thermal relaxation channel from Appendix A). The results for noisy simulation can be found in

	acetone	propen-2-ol
Full system CCSD energies relative to acetone	0	24.070
Exact active-space system energies relative to acetone	0	23.773
active-space system VQE energies relative to acetone	0	22.801
active-space system VQE Error	2.787	1.805

TABLE IV: Energy calculations of acetone and propen-2-ol in kcal/mol with the VQE results calculated on a noisy simulator. The energies are presented in the form of energy gaps (relative to acetone). Exact energies of the active-space systems were calculated by the diagonalization of the Hamiltonian. VQE energies were calculated using a 4-layer hardware-efficient ansatz (Fig. 4b) with Gaussian initialization for the parameters on a noisy simulator (Appendix A). VQE errors were calculated from the energy difference between VQE energies and the exact energies of the active-space systems.

Table IV where the VQE errors and the energy gaps are found to be close to chemical accuracy.

III. DISCUSSION

In this work, we have proposed a hybrid quantum-classical workflow for the prediction of the relative stability of tautomers. It is of the pharmaceutical industry's interest to accurately predict the preferred tautomeric state of a given molecules and its computation methodology. However, the computational requirements using classical approaches for quantum chemistry are too expensive, so we have made use of quantum simulation for electronic structure calculations.

We are aware of some studies that are making the effort to realize quantum simulation as a practical application for quantum chemistry. Recently, Tazhigulov et al. [38] simulated strongly-correlated molecules that are more relevant to real world problems on Google's Sycamore quantum processor. They mapped the electronic structures of iron-sulfur molecular clusters and α-RuCl$_3$ into low-energy spin models using the results from theoretical and spectroscopic studies. This simplification reduced the qubit requirements and rendered the possibility of simulating larger molecules. They then used the finite-temperature version of quantum imaginary time evolution to capture physical observables. However, the circuit depth for the evolution were too large for the quantum device. Thus, they recompiled the circuit on a classical noiseless emulator, and numerous error mitigation techniques were employed. One of their experiments that used 9 qubits and 82 two-qubit gates achieved acceptable accuracy, while the data for the experiment that used 11 qubits and 310 two-qubit gates were not quantitatively meaningful. This research benchmarked the performance and limitation of quantum simulation on quantum hardware. It is, therefore, important to reduce hardware requirements of physical simulation to this limitation (e.g. using similar qubit and two-qubit gate counts).

Instead of utilizing empirical information, we have used *ab initio* quantum chemistry methods to reduce quantum hardware requirements. Besides, the problem of interest is to find the relative ground state energy differences rather than absolute energies or other observables, so there are more rooms for orbital reduction. First, we optimize the geometry using B3LYP. Next, we use the natural orbitals from MP2 to select active orbitals. The selection criteria includes the occupation numbers and hardware limitation. To exploit as much quantum resources, we map the reduced Hamiltonian onto fewer qubits using QEE. Finally, we simulate the molecules/tautomers with 12 active spin-orbitals using VQE with 8 qubits and 80 two-qubit gates, which is fairly close to the

quantum hardware limitation from [38] (9 qubits and 82 two-qubit gates). In this work, not much efforts were done on exploring the best ansatz circuit for the VQE simulation, but we have found that an alternative ansatz circuit with 28 two-qubit gates can sufficiently provide accurate VQE calculations for the acetone system. This two-qubit gate count is much fewer than the quantum hardware limitation from [38] which shows the applicability of our methodology.

The VQE results for predicted stability agrees with those from full system CCSD results. For the prediction of acetone, its keto form is more stable than its enol form. For the case of Edaravone, its keto form is the most stable and its amine form is the most unstable. Note that these results did not incorporate solvent effects, which is an important factor in consideration of tautomerism in biological systems. Continuum solvation models, such as Polarizable Continuum Model (PCM)[62], the Solvation Model based on Density (SMD)[63] or the Conductor-like Screening Model (COSMO)[64], can be used to take into account the solvent effects. For example the PCM-VQE algorithm provides a self-consistent way to include the solvent effect which does not require extra quantum resources.

The CCSD results are considered to be the solution for the comparison with VQE results since FCI is too expensive for such large systems. Nevertheless, there could be even larger systems where CCSD starts becoming intractable. This is where our hybrid quantum-classical workflow could come into play because MP2 (which provides natural orbital occupancy) is computationally cheaper than CCSD and quantum simulation has the potential of becoming a more advantageous approach.

Note that even though the VQE data are from noiseless simulator, there are still some VQE errors. The reasons could be that either the ansatz circuit could not capture the numerically true ground states or the parameter optimization stopped at local minimum or got stuck on barren plateau. Nevertheless, the errors are within chemical accuracy and does not affect the stability prediction.

The inclusion of active space selection process in this work alleviates the burden on quantum resources. Of course a quantum device with more qubits that are less error-prone can bring value into quantum chemistry, but we are not close to achieving large-scale fault-tolerant in any architecture. Therefore, the workflow and methodology for quantum chemistry presented here is valuable in the NISQ era. As quantum computers will certainly be improved, the active orbital selection criteria can be adjusted such that more orbitals can be included. Such versatility stems from the fact that the limitation of quantum hardware is included in the selection criteria.

Given the importance of tautomerism in pharmaceutical research, our work aims to extend the applications of quantum computing in drug discovery and predictions of chemical events. To realize useful applications for quantum computers in near term, hybrid methods and reasonable classical preprocessing are necessary. On the other hand, the classical methods should be adaptable to the advancement of future quantum computers. Certain modifications on the quantum subroutines can also be done to improve the whole picture. We have employed QEE to increase the maximal number of spin-orbitals that can be simulated on limited quantum resources. Besides, to avoid the local minimum/barren plateau problem seen in this work, one could employ some parameter optimization strategies such as the initialization strategy proposed in [61] and [65] or one could change the VQE subroutine to methods such as imaginary time evolution [66] for better convergence when future hardware admits.

[1] L. Antonov, *Tautomerism : Concepts and Applications in Science and Technology* (Wiley, 2016).
[2] P. Muller, J. Macromol. Sci. Part A Pure Appl. Chem. **66**, 1077 (1994).

[3] I. Alkorta, P. Goya, J. Elguero, and S. P. Singh, National Academy Science Letters **30**, 139 (2007).

[4] W. Wang, H. W. Hellinga, and L. S. Beese, Proceedings of the National Academy of Sciences **108**, 17644 (2011).

[5] K. Bebenek, L. C. Pedersen, and T. A. Kunkel, Proceedings of the National Academy of Sciences **108**, 1862 (2011).

[6] Y. C. Martin, J. Comput. Aided Mol. Des. **23**, 693 (2009).

[7] W. A. Warr, J. Comput. Aided Mol. Des. **24**, 497 (2010).

[8] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).

[9] C. Dykstra, G. Frenking, and K. Kim, *Theory and Applications of Computational Chemistry: the First Forty Years.* (Elsevier Science, Amsterdam, 2011).

[10] R. P. Feynman, Int J Theor Phys **21**, 467 (1982).

[11] P. Jordan and E. Wigner, Z. Physik **47**, 631 (1928).

[12] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **65**, 042323 (2002).

[13] A. Y. Kitaev, arXiv:quant-ph/9511026 (1995), arXiv: quant-ph/9511026.

[14] J. Du, N. Xu, X. Peng, P. Wang, S. Wu, and D. Lu, Phys. Rev. Lett. **104**, 030502 (2010).

[15] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, Nature Chem **2**, 106 (2010).

[16] Z. Li, M.-H. Yung, H. Chen, D. Lu, J. D. Whitfield, X. Peng, A. Aspuru-Guzik, and J. Du, Sci Rep **1**, 88 (2011).

[17] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Phys. Rev. X **6**, 031007 (2016).

[18] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, Phys. Rev. Lett. **118**, 100503 (2017).

[19] R. Santagati, J. Wang, A. A. Gentile, S. Paesani, N. Wiebe, J. R. McClean, S. Morley-Short, P. J. Shadbolt, D. Bonneau, J. W. Silverstone, D. P. Tew, X. Zhou, J. L. O'Brien, and M. G. Thompson, Sci. Adv. **4**, eaap9646 (2018).

[20] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, and J. Wrachtrup, ACS Nano **9**, 7769 (2015).

[21] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **83**, 5162 (1999).

[22] A. Aspuru-Guzik, Science **309**, 1704 (2005).

[23] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat Commun **5**, 4213 (2014).

[24] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New J. Phys. **18**, 023023 (2016).

[25] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).

[26] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature **567**, 491 (2019).

[27] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo, Phys. Rev. Applied **9**, 044036 (2018).

[28] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Phys. Rev. Lett. **120**, 110501 (2018).

[29] D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, Phys. Rev. A **92**, 062318 (2015).

[30] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe, Phys. Rev. A **91**, 022311 (2015).

[31] K. Sugisaki, S. Yamamoto, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui, J. Phys. Chem. A **120**, 6459 (2016).

[32] K. Sugisaki, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui, ACS Cent. Sci. **5**, 167 (2019).

[33] Y. Du, M.-H. Hsieh, T. Liu, S. You, and D. Tao, PRX Quantum **3**, 030901 (2022).

[34] Google AI Quantum and Collaborators*†, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen,

Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. Dunsworth, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, H. Neven, M. Y. Niu, T. E. O'Brien, E. Ostby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, M. Szalay, T. Y. Takeshita, A. Vainsencher, T. White, N. Wiebe, Z. J. Yao, P. Yeh, and A. Zalcman, Science **369**, 1084 (2020).

[35] Y. Nam, J.-S. Chen, N. C. Pisenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf, K. M. Beck, A. Blinov, V. Chaplin, M. Chmielewski, C. Collins, S. Debnath, K. M. Hudek, A. M. Ducore, M. Keesan, S. M. Kreikemeier, J. Mizrahi, P. Solomon, M. Williams, J. D. Wong-Campos, D. Moehring, C. Monroe, and J. Kim, npj Quantum Inf **6**, 33 (2020).

[36] C. Cao, J. Hu, W. Zhang, X. Xu, D. Chen, F. Yu, J. Li, H. Hu, D. Lv, and M.-H. Yung, arXiv:2109.02110 [quant-ph] (2021), arXiv: 2109.02110.

[37] W. Li, Z. Huang, C. Cao, Y. Huang, Z. Shuai, X. Sun, J. Sun, X. Yuan, and D. Lv, arXiv:2109.08062 [quant-ph] (2021), arXiv: 2109.08062.

[38] R. N. Tazhigulov, S.-N. Sun, R. Haghshenas, H. Zhai, A. T. K. Tan, N. C. Rubin, R. Babbush, A. J. Minnich, and G. K.-L. Chan, arXiv:2203.15291 [quant-ph] (2022), arXiv: 2203.15291.

[39] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, arXiv:1701.08213 [quant-ph] (2017), arXiv: 1701.08213.

[40] N. Moll, A. Fuhrer, P. Staar, and I. Tavernelli, J. Phys. A: Math. Theor. **49**, 295301 (2016).

[41] R. Babbush, D. W. Berry, Y. R. Sanders, I. D. Kivlichan, A. Scherer, A. Y. Wei, P. J. Love, and A. Aspuru-Guzik, Quantum Sci. Technol. **3**, 015006 (2018), 1506.01029.

[42] M. Steudtner and S. Wehner, New J. Phys. **20**, 063010 (2018).

[43] W. Kirby, B. Fuller, C. Hadfield, and A. Mezzacapo, arXiv:2109.14465 [quant-ph] (2021), 2109.14465.

[44] Y. Shee, P.-K. Tsai, C.-L. Hong, H.-C. Cheng, and H.-S. Goan, arXiv:2110.04112 [cond-mat, physics:physics, physics:quant-ph] (2021), arXiv: 2110.04112.

[45] P. C. Rathi, R. F. Ludlow, and M. L. Verdonk, J. Med. Chem. **63**, 8778 (2020).

[46] K. Watanabe, M. Tanaka, S. Yuki, M. Hirai, and Y. Yamamoto, J. Clin. Biochem. Nutr. **62**, 20 (2018).

[47] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison, J. Cheminform. **4**, 17 (2012).

[48] T. A. Halgren, J. Comput. Chem. **17**, 616 (1996).

[49] T. A. Halgren, J. Comput. Chem. **17**, 520 (1996).

[50] T. A. Halgren, J. Comput. Chem. **17**, 553 (1996).

[51] T. A. Halgren, J. Comput. Chem. **17**, 490 (1996).

[52] T. A. Halgren and R. B. Nachbar, J. Comput. Chem. **17**, 587 (1996).

[53] A. D. Becke, J. Chem. Phys **98**, 5648 (1993).

[54] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, The Journal of physical chemistry **98**, 11623 (1994).

[55] M. Head-Gordon, J. A. Pople, and M. J. Frisch, Chem. Phys. Lett. **153**, 503 (1988).

[56] Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, Phys. Rev. Research **2**, 033125 (2020).

[57] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, Adv Quantum Tech **2**, 1900070 (2019).

[58] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nat Commun **9**, 4812 (2018).

[59] K. Zhang, M.-H. Hsieh, L. Liu, and D. Tao, Toward trainability of quantum neural networks (2020).

[60] K. Zhang, M.-H. Hsieh, L. Liu, and D. Tao, Toward trainability of deep quantum neural networks (2021).

[61] K. Zhang, M.-H. Hsieh, L. Liu, and D. Tao, Gaussian initializations help deep variational quantum circuits escape from the barren plateau (2022).

[62] C. Amovilli, V. Barone, R. Cammi, E. Cancès, M. Cossi, B. Mennucci, C. S. Pomelli, and J. Tomasi, in *Advances in Quantum Chemistry*, Vol. 32, edited by P.-O. Löwdin (Academic Press, 1998) pp. 227–261.

[63] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, The Journal of Physical Chemistry B **113**, 6378 (2009).

[64] A. Klamt and G. Schüürmann, J. Chem. Soc. Perkin Trans. 2 , 799 (1993).
[65] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, Quantum **3**, 10.22331/q-2019-12-09-214 (2019).
[66] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, npj Quantum Inf **5**, 75 (2019).

Appendix A: Noise Model

For the noisy simulation of VQE in this work, we implemented the 8-qubit noise model using the Qiskit package. We have employed the thermal relaxation errors for the simulation where the relaxation time constants are shown in Table V and Table VI and the gate time and measurement time are shown in Table VII.

TABLE V: Relaxation time constants of the 8-qubit noise model used for noisy VQE simulation of acetone.

	Qubit 0	Qubit 1	Qubit 2	Qubit 3	Qubit 4	Qubit 5	Qubit 6	Qubit 7
T_1 (ms)	194.24	230.20	218.81	212.36	189.67	198.52	230.69	153.56
T_2 (ms)	188.28	219.84	226.43	165.82	236.03	204.83	196.17	166.16

TABLE VI: Relaxation time constants of the 8-qubit noise model used for noisy VQE simulation of propen-2-ol.

	Qubit 0	Qubit 1	Qubit 2	Qubit 3	Qubit 4	Qubit 5	Qubit 6	Qubit 7
T_1 (ms)	134.45	147.61	208.15	211.67	170.22	208.24	200.66	242.95
T_2 (ms)	60.85	234.66	178.45	210.15	224.01	178.00	105.58	136.21

TABLE VII: Gate time and measurement time in nanoseconds of the 8-qubit noise models used for noisy VQE simulation of acetone and propen-2-ol.

U_1 gate	U_2 gate	U_3 gate	CNOT gate	Reset time	Measurement time
0	50	100	200	1000	1000

Appendix B: Natural Orbital Occupancy

Natural orbital occupancies were used to select active molecule orbital (MOs) sets in this work. The natural occupancy reflects the occupation state of a given orbital. With the natural orbital occupation numbers, we reduce the MO sets by removing those orbitals that are nearly virtual (natural orbital occupancy close to 0) or fully occupied (natural orbital occupancy close to 2), which are not likely involved in the tautomerization process.

Table VIII and Table IX show the MO energies and the natural orbital occupancies of acetone and Edaravone systems using MP2 calculations.

TABLE VIII: The eigenvalues of the MOs and the occupancies of the natural orbitals of the acetone tautomers calculated at the MP2/STO-3G level.

MO	Keto		Enol	
	Eigenvalue (Hatree)	Occupancy	Eigenvalue (Hatree)	Occupancy
0	-20.26834	2.000004	-20.28742	2.000003
1	-11.12074	2.000002	-11.08992	2.000002
2	-11.04974	2.000000	-11.05175	2.000000
3	-11.04853	2.000000	-10.98477	2.000000
4	-1.32547	1.997842	-1.32047	1.998189
5	-0.98289	1.993911	-0.99634	1.995010
6	-0.91791	1.992408	-0.89618	1.992262
7	-0.68101	1.991211	-0.71645	1.990954
8	-0.59048	1.987791	-0.61208	1.988330
9	-0.57718	1.987463	-0.58038	1.987414
10	-0.57548	1.987179	-0.55659	1.987216
11	-0.52423	1.987075	-0.50819	1.987117
12	-0.49429	1.986979	-0.46796	1.985933
13	-0.48573	1.983741	-0.45614	1.982110
14	-0.39341	1.978252	-0.43099	1.979630
15	-0.32185	1.942575	-0.26688	1.950616
16	0.29921	0.061559	0.33083	0.054016
17	0.61656	0.021915	0.56090	0.020035
18	0.66220	0.016135	0.61885	0.016292
19	0.67844	0.015970	0.69870	0.015365
20	0.68052	0.012180	0.71459	0.013645
21	0.71925	0.011860	0.71780	0.012069
22	0.73087	0.011671	0.72110	0.011897
23	0.73720	0.011361	0.75222	0.011141
24	0.81362	0.010535	0.93500	0.010535
25	1.02163	0.010382	1.04183	0.010220

Appendix C: Selective combination of active molecular Orbitals

A selection of active molecular orbital sets were chosen and their corresponding ground state energies were calculated at the level of CCSD theory for all tautomers of the two example.

Table X and Table XI show the relative CCSD energies for several active-space MO of the acetone and Edaravone systems.

(a) VQE energy errors (Gaussian Init.).

(b) VQE energies (Gaussian Init.).

(c) VQE energy errors (HF Init.).

(d) VQE energies (HF Init.).

FIG. 5: The relationship between acetone VQE results and the number of hardware-efficient ansatz layers. (a) and (b) are the results when Gaussian-initialized parameters are used while (c) and (d) are the results when initial states are the HF states. All the x-axes represent the number of hardware-efficient circuit layers. For (a) and (c), the y-axes represent VQE energy errors of the keto or enol form of acetone, and the dashed lines represent chemical accuracy (Chem. Accu.). For (b) and (d), the y-axes represent the energies relative to the keto form (so VQE energies, CCSD energies, and exact energies of the keto form are set to zero). Note that the exact energies (Exact in the legend) are the results from exact diagonalization of the active-space Hamiltonian for each tautomer, and CCSD full in the legend represents the CCSD energy of the full system relative to the keto form.

(a) VQE energy errors (Gaussian Init.).

(b) VQE energies (Gaussian Init.).

(c) VQE energy errors (HF Init.).
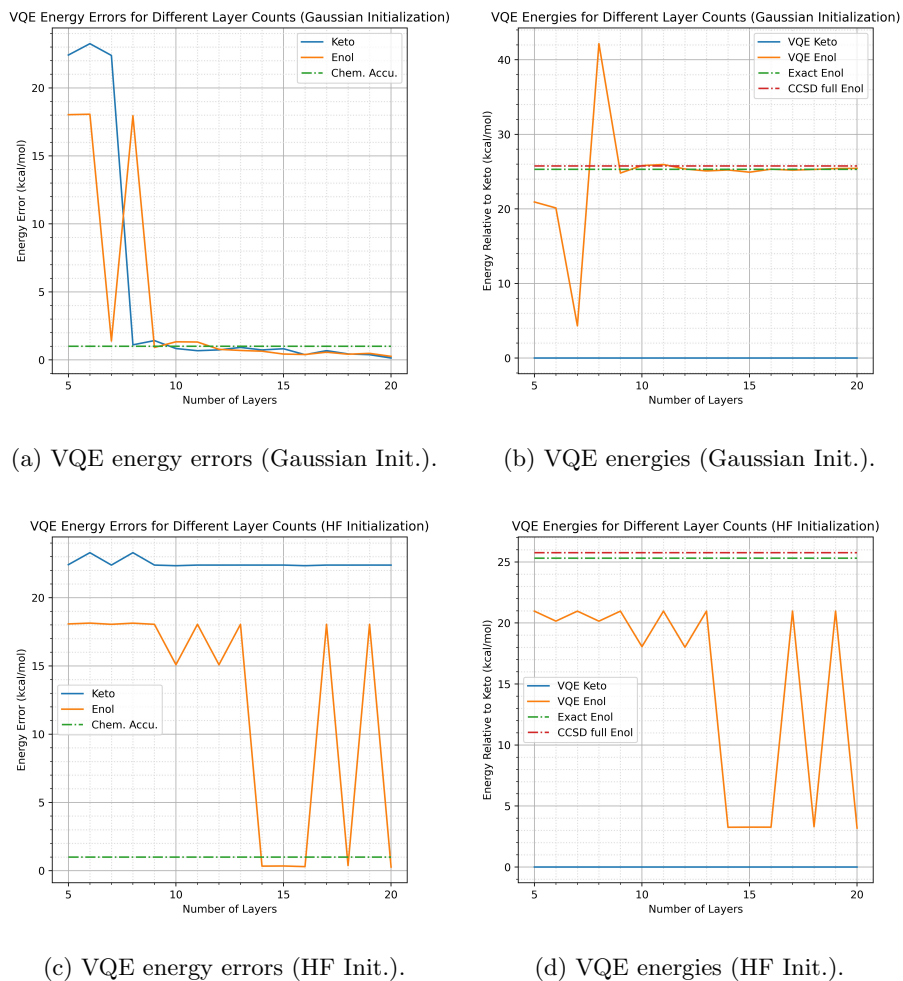

(d) VQE energies (HF Init.).

FIG. 6: The relationship between Edaravone VQE results and the number of hardware-efficient ansatz layers. (a) and (b) are the results when Gaussian-initialized parameters are used while (c) and (d) are the results when initial states are the HF states. All the x-axes represent the number of hardware-efficient circuit layers. For (a) and (c), the y-axes represent VQE energy errors of the keto, enol, or amine form of Edaravone, and the dashed lines represent chemical accuracy (Chem. Accu.). For (b) and (d), the y-axes represent the energies relative to the keto form (so VQE energies, CCSD energies, and exact energies of the keto form are set to zero). Note that the exact energies (Exact in the legend) are the results from exact diagonalization of the active-space Hamiltonian for each tautomer, and CCSD full in the legend represents the CCSD energies of the full systems relative to the keto form.
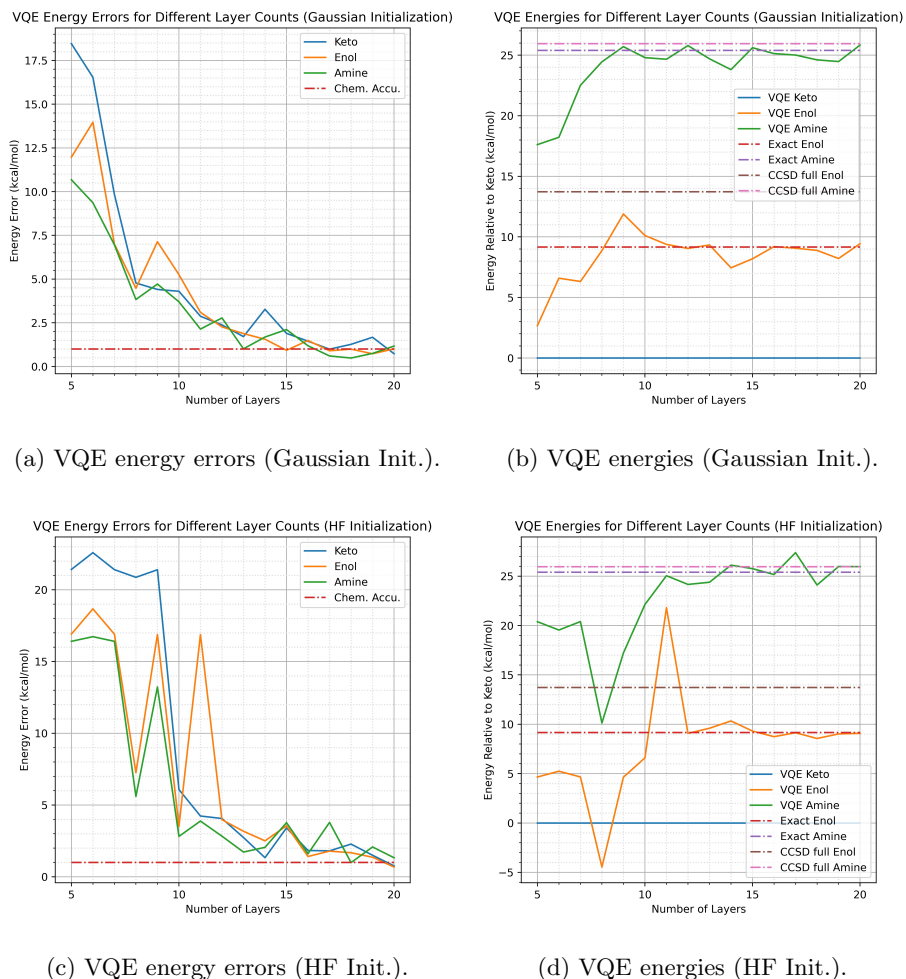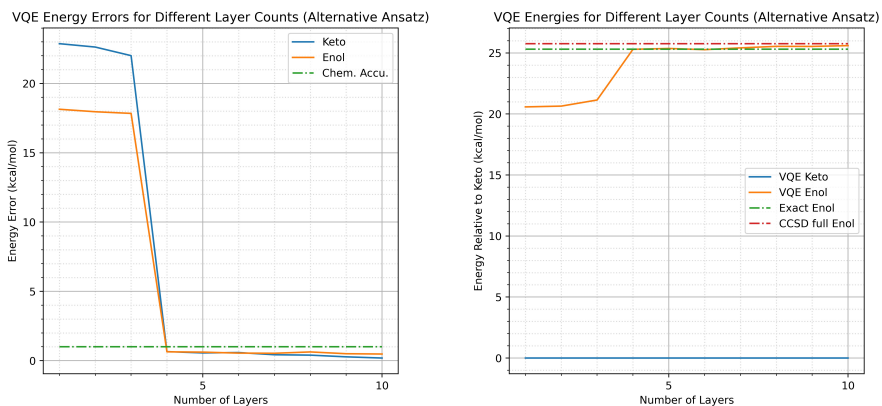
(a) VQE energy errors (Gaussian Init.).   (b) VQE energies (Gaussian Init.).

FIG. 7: The relationship between acetone VQE results and the number of the alternative ansatz layers from Fig. 4b. Gaussian-initialized parameters are used. All the x-axes represent the number of hardware-efficient circuit layers. For (a), the y-axis represents VQE energy errors of the keto or enol form of acetone, and the dashed line represents chemical accuracy (Chem. Accu.). For (b), the y-axis represents the energies relative to the keto form (so VQE energies, CCSD energies, and exact energies of the keto form are set to zero). Note that the exact energies (Exact in the legend) are the results from exact diagonalization of the active-space Hamiltonian for each tautomer, and CCSD full in the legend represents the CCSD energy of the full system relative to the keto form.

TABLE IX: The eigenvalues of the MOs and the occupancies of the natural orbitals of the Edaravone tautomers calculated at the MP2/STO-3G level.

	Keto		Enol		Amine	
MO	Eigenvalue	Occupancy	Eigenvalue	Occupancy	Eigenvalue	Occupancy
0	-20.26132	2.000007	-20.3286	2.000007	-20.21207	2.000004
1	-15.41341	2.000004	-15.4416	2.000002	-15.41433	2.000002
2	-15.39275	2.000002	-15.3295	2.000002	-15.39921	2.000002
3	-11.17214	2.000002	-11.13800	2.000002	-11.14758	2.000002
4	-11.09193	2.000002	-11.0931	2.000002	-11.10969	2.000001
5	-11.08892	2.000001	-11.0503	2.000001	-11.08820	2.000001
6	-11.08030	2.000001	-11.0441	2.000001	-11.08076	2.000001
7	-11.07000	2.000001	-11.0357	2.000001	-11.03183	2.000001
8	-11.02927	2.000001	-11.0351	2.000001	-11.03030	2.000001
9	-11.02916	2.000000	-11.03500	2.000001	-11.02411	2.000000
10	-11.02174	2.000000	-11.0335	2.000000	-11.02210	2.000000
11	-11.02090	2.000000	-11.0315	2.000000	-11.02121	2.000000
12	-11.02028	2.000000	-10.9921	2.000000	-11.01432	2.000000
13	-1.35162	1.998050	-1.36683	1.998025	-1.32861	1.998176
14	-1.28962	1.995627	-1.29727	1.995976	-1.27162	1.994969
15	-1.11989	1.993057	-1.11338	1.994566	-1.11756	1.993101
16	-1.06485	1.992505	-1.05489	1.992798	-1.06206	1.992443
17	-0.99215	1.992335	-0.98405	1.992418	-1.00022	1.992341
18	-0.96551	1.991838	-0.96177	1.991945	-0.96884	1.991857
19	-0.95409	1.990887	-0.95266	1.990953	-0.95565	1.990873
20	-0.91721	1.990643	-0.90150	1.990852	-0.91324	1.990719
21	-0.79819	1.989859	-0.79539	1.989632	-0.78523	1.990324
22	-0.77423	1.988589	-0.77676	1.988703	-0.77646	1.989263
23	-0.72782	1.987920	-0.73127	1.988055	-0.73674	1.988787
24	-0.69806	1.987380	-0.69621	1.987876	-0.71001	1.988546
25	-0.64911	1.987330	-0.68063	1.987247	-0.66793	1.987863
26	-0.62823	1.987190	-0.63355	1.987126	-0.63605	1.987103
27	-0.62162	1.986792	-0.60903	1.986818	-0.61366	1.986961
28	-0.58933	1.986318	-0.59742	1.986296	-0.58833	1.986617
29	-0.56595	1.985577	-0.57066	1.985343	-0.56420	1.986337
30	-0.55252	1.985491	-0.55774	1.983740	-0.56325	1.985574
31	-0.54924	1.983691	-0.54279	1.983326	-0.55100	1.983833
32	-0.53899	1.983332	-0.53714	1.983290	-0.53788	1.983331
33	-0.52916	1.983207	-0.52954	1.981034	-0.53218	1.983047
34	-0.52595	1.980687	-0.51471	1.980603	-0.52911	1.980885
35	-0.51468	1.979081	-0.49153	1.979242	-0.50496	1.979233
36	-0.49568	1.978694	-0.47726	1.978879	-0.48472	1.978964
37	-0.45479	1.978138	-0.46014	1.977825	-0.45968	1.978061
38	-0.44884	1.977486	-0.44551	1.977524	-0.45128	1.976938
39	-0.43422	1.975835	-0.44170	1.974787	-0.43474	1.975575
40	-0.41856	1.972720	-0.43297	1.971615	-0.42470	1.974069
41	-0.39887	1.968219	-0.35506	1.968096	-0.35961	1.968380
42	-0.32380	1.947555	-0.33103	1.958506	-0.31246	1.957051
43	-0.32217	1.941835	-0.28577	1.941071	-0.28196	1.939416
44	-0.28022	1.938252	-0.24767	1.937216	-0.26851	1.938023
45	-0.22348	1.933366	-0.24342	1.934966	-0.21386	1.934162
46	0.25533	0.068063	0.23986	0.073750	0.23389	0.071535
47	0.25768	0.065302	0.26492	0.064523	0.26903	0.065091
48	0.27132	0.064662	0.32588	0.064330	0.28245	0.062220
49	0.35220	0.059913	0.38423	0.051273	0.40600	0.048977
50	0.51155	0.032174	0.49939	0.031852	0.50366	0.032236
51	0.51424	0.026326	0.53225	0.026128	0.52337	0.025914
52	0.54395	0.023414	0.53709	0.022501	0.57065	0.022931
53	0.59018	0.021915	0.60062	0.021388	0.57568	0.022116
54	0.61087	0.021247	0.60867	0.020487	0.61116	0.020147
55	0.65256	0.019565	0.64232	0.019322	0.62894	0.018357
56	0.67183	0.018364	0.66297	0.018339	0.65775	0.018193
57	0.67780	0.018049	0.68513	0.018043	0.66583	0.018053
58	0.67839	0.017572	0.69013	0.017615	0.68247	0.017503
59	0.68719	0.017058	0.70808	0.016812	0.69122	0.016896
60	0.71836	0.016353	0.72472	0.015515	0.70617	0.016056
61	0.72364	0.015473	0.72612	0.015122	0.71606	0.015525
62	0.72700	0.014330	0.73009	0.014612	0.73061	0.014222
63	0.73249	0.012992	0.74037	0.014018	0.73815	0.013586
64	0.73670	0.012567	0.76168	0.013661	0.76211	0.013127
65	0.77141	0.012296	0.81055	0.012333	0.77341	0.012328
66	0.85449	0.012056	0.85656	0.011997	0.87823	0.012037
67	0.87814	0.012016	0.88338	0.011977	0.88962	0.011974
68	0.89312	0.011976	0.88839	0.011952	0.89827	0.011929
69	0.90300	0.011815	0.90412	0.011807	0.90475	0.011797
70	0.93342	0.011758	0.94604	0.011744	0.92936	0.011747
71	0.98338	0.011587	1.00630	0.011687	0.97650	0.011703
72	1.04928	0.011067	1.08665	0.010419	1.09233	0.010538
73	1.09713	0.010395	1.11121	0.010313	1.12981	0.010254
74	1.15351	0.010190	1.14982	0.010110	1.15709	0.010169

TABLE X: Active space selection of molecular orbitals for acetone. Relative ground state energies of the keto, and enol forms of acetone were calculated for selected combinations of molecular orbitals. Listed ground state energies are relative to the keto form. The values are shown in kcal/mol.

Active MO	# of MO	# of SO	# of e	Keto	Enol
0-25	26	52	32	0	24.070
10-19	10	20	12	0	18.716
11-20	10	20	10	0	22.478
11-19	9	18	10	0	21.260
12-20	9	18	8	0	21.374
12-19	8	16	8	0	19.888
13-19	7	14	6	0	20.325
13-18	6	12	6	0	20.565
14-19	6	12	4	0	23.766

TABLE XI: Active space selection of molecular orbitals for Edaravone. Relative ground state energies of the keto, enol and amine forms of Edaravone were calculated for selected combinations of molecular orbitals. Listed ground state energies are relative to the keto form. The values are shown in kcal/mol.

Active MO	# of MO	# of SO	# of e	Keto	Enol	Amine
0-74	75	150	92	0	13.726	25.947
41-50	10	20	10	0	-12.793	4.904
41-49	9	18	10	0	-12.959	6.556
42-50	9	18	8	0	-12.650	23.293
42-49	8	16	8	0	-12.874	22.420
43-49	7	14	6	0	8.583	23.892
43-48	6	12	6	0	5.667	18.618
44-49	6	12	4	0	9.197	25.398