

Nuclear-Electronic Orbital and Nonadiabatic Dynamics Calculations on Molecular Systems

PI: Sharon Hammes-Schiffer, Blue Waters Professor, Chemistry Department, University of Illinois at Urbana-Champaign

Collaborators: Kurt Brorsen, Puja Goyal, and Alexander Soudackov, Chemistry Department, University of Illinois at Urbana-Champaign

Corresponding author: Sharon Hammes-Schiffer, shs3@illinois.edu

Executive Summary

The quantum mechanical behavior of nuclei plays an important role in a wide range of chemical and biological processes. The inclusion of nuclear quantum effects and nonadiabatic effects between nuclei and electrons in computer simulations is challenging. Our group has developed the nuclear-electronic orbital (NEO) method for treating electrons and select nuclei quantum mechanically on the same level using an orbital-based formalism. Our group has also developed mixed quantum mechanical/molecular mechanical nonadiabatic dynamics methods for simulating photoinduced proton-coupled electron transfer processes. Both types of calculations require a large number of processors and a substantial amount of memory and storage. The highly parallel computing system on Blue Waters is advantageous for these codes because it provides the capability of splitting the large number of calculations and storage requirements over many nodes with efficient inter-node communication. We request 240,000 node hours divided equally between the NEO project and the nonadiabatic dynamics project.

1. Description of Research Activities and Results

A. Key Challenges

The overall objective of our research on Blue Waters is the development and application of computational methods for investigating proton-coupled electron transfer (PCET) reactions, which are important in a variety of biological and chemical processes. The proposed research on Blue Waters will focus on two different computational approaches developed by my group: (1) the incorporation of non-Born-Oppenheimer effects between electrons and protons into electronic structure calculations using the nuclear-electronic orbital (NEO) method to elucidate the fundamental principles of PCET; and (2) the simulation of photoinduced PCET processes to guide the design of more effective solar devices for energy production and storage.

The first key challenge that is being addressed is the development of computational methods that include both nuclear quantum effects and nonadiabatic effects for molecular systems. The inclusion of nuclear quantum effects such as zero point energy and tunneling in electronic structure calculations is important for the study of a variety of chemical systems, particularly those involving hydrogen transfer or hydrogen-bonding interactions. Moreover, nonadiabatic effects, also denoted non-Born-Oppenheimer effects, between electrons and certain nuclei are significant for many of these systems. In this case, the electrons cannot be assumed to respond instantaneously to the nuclear motions, and the concept of the nuclei moving on a single electronic potential energy surface is no longer valid. This type of nonadiabaticity has been shown to play a critical role in PCET reactions. Although a large number of computational methods have been developed to study nuclear quantum dynamics, most of these methods invoke the Born-Oppenheimer approximation between electrons and nuclei and therefore neglect these important non-Born-Oppenheimer effects. The few electronic structure methods that include

non-Born-Oppenheimer effects are extremely computationally expensive and therefore cannot be applied to systems with more than only a few electrons and nuclei.

Our objective is to develop computationally tractable methods that include non-Born-Oppenheimer effects as well as nuclear quantum effects and to apply these methods to chemically and biologically important systems. Specifically, we are developing the nuclear-electronic orbital (NEO) method, which treats electrons and select nuclei quantum mechanically on the same level using an orbital-based formalism. The NEO method is ideal for studying PCET because the timescale for proton tunneling is often faster than the timescale for electronic transitions, thereby leading to a breakdown of the Born-Oppenheimer approximation. In applications of the NEO method to PCET, all electrons and one or a few protons are treated quantum mechanically, and a mixed nuclear-electronic time-independent Schrödinger equation is solved using explicitly correlated wavefunctions.

The second key challenge that is being addressed is the development of computational methods for simulating the nonadiabatic dynamics of photoinduced PCET processes for molecules in solution. These approaches differ from the NEO method in that they probe the time-dependent, nonequilibrium dynamics following photoexcitation of PCET systems in solution. Although many computational methods have been developed to study photoinduced electron transfer processes, typically these methods treat all nuclei classically. In PCET processes, however, the quantum mechanical behavior of the transferring proton has been shown to be critical. Thus, our objective is to develop computationally tractable methods for simulating the nonadiabatic dynamics of condensed phase systems while including the essential nuclear quantum mechanical effects. In our approach, the transferring proton is treated quantum mechanically with Fourier grid-based methods, the solvation effects are included with explicit solvent molecules, the potential energy surfaces are generated on-the-fly with multireference electronic structure methods, and the nonadiabatic dynamics is described with a surface hopping algorithm. This methodology enables simulations that will elucidate the roles of solute and solvent dynamics, vibrational relaxation, proton delocalization, and electron-proton coupling, as well as identify the physical properties of PCET systems that determine the relaxation timescales.

B. Why it Matters

Proton-coupled electron transfer (PCET) reactions play a vital role in a broad range of biological and chemical processes. The coupling between electrons and protons is central to photosynthesis, respiration, and enzyme reactions, as well as solar cells and other energy conversion devices. Understanding the fundamental physical principles underlying PCET reactions will have broad implications ranging from drug design to the design of more effective catalysts for solar energy devices.

The NEO method will have applications beyond PCET systems. It will be useful for studying hydrogen-bonding interactions, hydrogen tunneling, and geometric isotope effects in a variety of chemical and biological systems. The NEO method may also have applications in the calculation and interpretation of nuclear magnetic resonance (NMR) spectra. It will also be useful for studying positronic systems, where the electrons and positrons are treated quantum mechanically. Positron emission tomography is an imaging technique that plays an important role in medical treatments. Understanding the fundamental behavior of positrons could lead to technological advances in this technique.

Nonadiabatic dynamics methods that include the nuclear quantum effects of the transferring proton will help guide the design of more effective catalysts for energy production and storage. Such methods will also have applications to biological systems such as photosynthesis and green fluorescent protein.

C. Why Blue Waters

The extraordinary capabilities of Blue Waters will advance both of these projects because of the highly parallelizable nature of our computer codes. The NEO code that will be used for the first project has been written in-house and has been extensively tested on several different computer architectures. Currently the computational bottleneck in a NEO calculation is the calculation of the multi-particle integrals. Since these integrals can be calculated completely independently from one another, this part of the calculation is embarrassingly parallelizable. Thus, we applied the OpenMP protocol to parallelize the calculation of these integrals, providing almost perfect scaling with respect to the number of threads. Recently we adopted a hybrid MPI/OpenMP protocol, which obviates the need for all integrals to be stored simultaneously and allows the division of the calculation over different nodes. This version of the code also demonstrates very good scalability with respect to the number of MPI processes. A supercomputer such as Blue Waters maximizes scalability. The large memory requirements of storing the multi-particle integrals render this calculation impossible on other computing resources, where a large number of nodes cannot be used simultaneously. The hybrid MPI/OpenMP implementation combined with Blue Waters provides the capability of splitting the large number of calculations and storage requirements over many nodes. As our code has demonstrated excellent scaling, we directly benefit from using a large number of nodes on Blue Waters simultaneously with very little overhead.

The code that will be used for the photoinduced PCET simulations in the second project is an in-house modified MOPAC code and is also highly parallelizable. For each system, 100–1000 molecular dynamics trajectories must be propagated. These trajectories are computationally expensive because the excited state potential energy surfaces are generated on-the-fly with electronic structure methods. Within each trajectory, the computational bottleneck is the calculation of the electronic potential energies, forces, and nonadiabatic couplings on a grid along the transferring proton coordinate at each molecular dynamics time step. However, each of these grid point calculations can be carried out independently. After these quantities are collected on the grid, the calculation of the final proton vibrational wavefunction is computationally inexpensive. At each time step, the typical number of electronic structure calculations is 96 (the product of 32 grid points and 3 electronic states), which can all be performed simultaneously if distributed over three Blue Waters nodes using MPI parallelization. At the end of these independent calculations, all processors can relay the necessary information to the master processor for computation of the proton vibrational wavefunction. The large number of nodes on Blue Waters and the ability to communicate between nodes efficiently through Cray Gemini torus interconnect are crucial for making these simulations computationally viable. The storage needs for these simulations will also benefit from Blue Waters because each trajectory will require a substantial amount of disk space to enable subsequent analysis of the ensemble of trajectories. The code for these simulations has already been tested and parallelized. The initial tests have been performed on our in-house computer cluster, but the extraordinary capabilities of Blue Waters will be necessary to provide a sufficient number of nodes with efficient inter-node communication and storage to allow these types of simulations for systems relevant to solar energy conversion.

D. Accomplishments

Over the past year, we have made advances with the NEO method that enabled us to perform the first NEO calculations on real molecules. Moreover, these advances will allow us to study even larger systems in the near future. In the NEO approach, typically all electrons and one or a few protons are treated quantum mechanically, and a mixed nuclear-electronic time-independent Schrödinger equation is solved. The nuclear and electronic

molecular orbitals in the nuclear-electronic wavefunction are expanded in terms of nuclear and electronic Gaussian basis sets, respectively, and application of the variational method leads to equations that must be solved iteratively to self-consistency using standard techniques. To include the essential electron-proton correlation, we developed an explicitly correlated method, denoted NEO-XCHF, which incorporates explicit electron-proton correlation with Gaussian-type geminal functions that depend on the electron-proton distance in the ansatz for the total nuclear-electronic wavefunction. The main computational expense arises from the two-, three-, and four-particle integrals that must be calculated and stored in memory or on disk. Although these explicitly correlated methods have been shown to be highly accurate for model systems, they are computationally expensive and are currently intractable for larger systems of chemical interest.

To improve the computational tractability, we developed a reduced version of NEO-XCHF, denoted NEO-RXCHF. In this approach, only select electronic orbitals are explicitly correlated to the nuclear orbital(s), and certain exchange terms are approximated, thereby substantially decreasing the number of multi-particle integrals that must be calculated. Moreover, we also developed a restricted basis set method, which leads to even greater computational savings. The NEO-RXCHF code was written in-house and has been tested extensively on several different computer architectures. The computational bottleneck is still the calculation of the multi-particle integrals. This computational bottleneck has been reduced by parallelizing the calculation of all multi-particle integrals involving a geminal function using a hybrid MPI/OpenMP approach. The hybrid MPI/OpenMP implementation shows good scaling with respect to MPI processes and the number of threads. This development has allowed the NEO code to take advantage of the full potential of Blue Waters. Additionally, the NEO-RXCHF method has been extended to systems with an odd number of electrons.

We have used Blue Waters to perform NEO-RXCHF calculations on three molecular systems: neutral hydrogen cyanide (HCN), positively charged hydrogen cyanide (HCN⁺), and the fluorine bihalide ion (FHF⁻). In all cases, the proton and all electrons were treated quantum mechanically within the NEO framework. We analyzed the nuclear densities of the protons and compared them to highly accurate grid-based densities used as benchmarks. Our calculations illustrated that this approach can provide accurate descriptions of the protons that are treated quantum mechanically. We also tested new approximate methods that will enable the study of larger proton-containing systems. This work resulted in a publication submitted to the *Journal of Chemical Physics* and currently under revision.

Over the past year, we also developed the methodology for mixed quantum mechanical/molecular mechanical (QM/MM) nonadiabatic dynamics simulations of photoinduced PCET processes. The potential energy surface is generated on-the-fly with a semiempirical floating occupation molecular orbital complete active space configuration interaction method for the solute molecule and a molecular mechanical force field for the explicit solvent molecules. The nonadiabatic dynamics is described by the fewest switches surface hopping method. We applied this method to the hydrogen-bonded *p*-nitrophenylphenol-*t*-butylamine complex solvated in 1,2-dichloroethane. The calculated relaxation times following photoexcitation are qualitatively consistent with experimental measurements. Overall, the calculations augment the original interpretation of the experimental data by providing evidence of proton transfer on the excited electronic state prior to decay to the ground state. The fundamental insights obtained from these simulations are also relevant to other photoinduced PCET processes. This work resulted in a publication that has been accepted for publication in *The Journal of Physical Chemistry B* and is posted on the ASAP web site of this journal. These simulations treated the proton classically and therefore did not require substantial computational resources.

Recently we have extended this QM/MM methodology to treat the proton quantum mechanically using Fourier grid methods. In this approach, the QM calculations of the energies, forces, and nonadiabatic derivative couplings must be performed on the one-dimensional grid along the proton coordinate, where the grid typically consists of 32 points. Thus, the quantum treatment of the proton is significantly more expensive and requires parallelization and capabilities of Blue Waters. We use the distributed data MPI strategy to parallelize these calculations, distributing the grid point calculations over different processors. Due to small communication overhead we expect excellent scaling for the parallel runs when the number of utilized processing elements (PEs) is equal to the product of the number of grid points and the number of included electronic states. We are currently using Blue Waters to apply this methodology to the same hydrogen-bonded complex studied with the classical treatment of the proton.

2. List of Publications and Presentations Associated with this Work

A. Publications that Featured Blue Waters Projects

- Sirjoosingh, M. V. Pak, K. Brorsen, and S. Hammes-Schiffer, "Quantum treatment of protons with the reduced explicitly correlated Hartree-Fock approach," *J. Chem. Phys.* (under revision).
- P. Goyal, C. A. Schwerdtfeger, A. V. Soudackov, and S. Hammes-Schiffer, "Nonadiabatic dynamics of photoinduced proton-coupled electron transfer in a solvated phenol-amine complex," *J. Phys. Chem. B* (in press, ASAP, DOI: 10.1021/jp5126969).

B. Presentations that Featured Blue Waters Projects

- Distinguished Lecture in Theoretical and Computational Chemistry, University of California at San Diego, San Diego, California, April 1, 2014 (invited talk): "Proton-Coupled Electron Transfer in Catalysis and Energy Conversion"
- Haines Lecture, University of South Dakota, April 7, 2014 (invited talk): "Proton-Coupled Electron Transfer in Energy Conversion Processes"
- Colloquium, University of Rochester, April 9, 2014 (invited talk): "Proton-Coupled Electron Transfer in Catalysis and Energy Conversion"
- Promoting Female Excellence in Theoretical and Computational Chemistry II, Oslo, Norway, June 13-15, 2014 (plenary talk): "Proton-Coupled Electron Transfer in Catalysis and Energy Conversion"
- PCET 2014: Second International Conference on Proton-Coupled Electron Transfer, Uppsala, Sweden, June 15-19, 2014 (invited talk): "Theoretical Perspectives of Proton-Coupled Electron Transfer and Applications to Catalysis"
- International Conference on Hydrogen Atom Transfer, Rome, Italy, June 22-26, 2014 (invited talk): "Theoretical Perspectives of Proton-Coupled Electron Transfer and Applications to Catalysis"
- American Conference of Theoretical Chemistry, Telluride, Colorado, July 20-25, 2014 (invited talk): "Theoretical Perspectives of Proton-Coupled Electron Transfer and Applications to Catalysis"
- American Chemical Society National Meeting, Symposium entitled Photoinduced Proton Transfer (PPT) in Chemistry and Biology, San Francisco, California, August 10-14, 2014 (invited talk converted to contributed): "Nonadiabatic Dynamics of Photoinduced Proton-Coupled Electron Transfer Processes"

- Northwestern University, Evanston, Illinois, October 24, 2014 (colloquium): “Proton-Coupled Electron Transfer in Catalysis and Energy Conversion”
- Princeton Center for Theoretical Science, Workshop on Numerical Approaches to Nonadiabatic Dynamics, November 21-22, 2014 (invited talk): “Avoiding the Born-Oppenheimer separation between electrons and protons: Nuclear-electronic orbital techniques and nonadiabatic dynamics”

C. Plan for Next Year

We request 240,000 node hours for the second year of the Blue Waters Professorship. The estimated usage per quarter is 60,000 node hours, corresponding to continuous, steady use over this time. We expect to require ~20 TB of storage per year. This allocation request can be justified on the basis of the projects described above. Although the NEO project greatly benefited from Blue Waters, we did not use our entire allocation during the first year because we were developing the code for both of these projects. The debugging process of the hybrid MPI/OpenMP was challenging and time consuming for these large, complex codes. The NEO code is now fully parallelized and tested, and we are ready for production calculations on larger molecular systems. The QM/MM nonadiabatic dynamics code is also parallelized and tested, and we are ready for production calculations with a quantized proton. Given that both codes are now parallelized, tested, and ready for production on Blue Waters, we expect to use our entire allocation this year. Currently several large jobs for the NEO project are in the queue on Blue Waters, and several jobs for the nonadiabatic dynamics project will be submitted over the next few weeks.

We plan to use 120,000 node hours for the first project. The explicitly correlated NEO calculation on the FHF^- molecule required over 1 TB of integrals that needed to be calculated and stored. The explicitly correlated NEO calculations on HCN and HCN^+ required a similar number of integrals to be stored. The two sets of integrals calculated over the past year used nearly one half of our allocation (120,000 node hours) even while taking advantage of various Blue Waters discounts for node hour charges. (The NEO calculations on HCN and HCN^+ only required the calculation of one set of integrals, therefore requiring two sets of integrals rather than three sets.) Over the next year, we plan to utilize Blue Waters to study two larger systems that are considered prototypes for two types of PCET reactions: the phenoxyl-phenol and the benzene-toluene self-exchange reactions. These systems will be computationally tractable due to several algorithmic and numerical improvements that have been implemented, but will still require ~60,000 node hours each. These applications will represent the first NEO calculations on chemically relevant PCET systems and will be a significant achievement.

The estimate of 60,000 node hours per system is obtained by assuming that the integrals for each system will require nine different calculations for the relevant configurations, each requiring 256 nodes for 24 hours. After the calculation of the integrals, the calculation of the electronic and nuclear densities is estimated to require 256 nodes for 18 hours. Table 1 provides a more detailed breakdown of the computational requirements for the NEO calculations. The NEO calculations on these two systems will require approximately 119,808 node hours. The calculation of the electronic and nuclear densities will require the most memory and will use ~20 GB of memory per node. The storage of the integrals for the phenoxyl-phenol and benzene-toluene calculations will require ~10 TB. We do not anticipate transferring any significant amount of data to or from Blue Waters for the NEO project.

Table 1. Computational requirements for the NEO calculations.

Number of integral calculations	18
Node hours per integral calculation ^a	6,144
Node hours for integral calculations	110,592
Number of density calculations	2
Node hours per density calculation	4,608
Node hours for density calculations	9,216
Total number of node hours	119,808

^aThe integral calculations will use 32 PEs per node.

We plan to use 120,000 node hours for the second project. We will study the hydrogen-bonded *p*-nitrophenylphenol-*t*-butylamine complex solvated in 1,2-dichloroethane that we studied previously with a classical treatment of the transferring proton. This application will represent the first nonequilibrium quantum dynamical study of photoinduced PCET. Moreover, these calculations will provide the kinetic isotope effect, which is the ratio of the relaxation time with protium and deuterium, as a prediction that can be tested experimentally.

Our estimate of the number of node hours is based on propagating 144 trajectories for each isotope, protium or deuterium, of the transferring hydrogen. For each isotope, the trajectories will differ in the initial coordinates, velocities, and occupied vibronic state. Each trajectory will be 5 ps long and will require ~417 node hours. Detailed information about the computational requirements for these QM/MM nonadiabatic dynamics simulations is provided in Table 2. The propagation of 288 nonadiabatic dynamics trajectories, including solvent effects and a quantum mechanical treatment of the proton, is estimated to require approximately 120,010 node hours. Each trajectory will require 10 GB of memory per node. Storage of these trajectories for subsequent analysis will require ~10 TB. We do not anticipate transferring any significant amount of data to or from Blue Waters for the QM/MM nonadiabatic dynamics project.

Table 2. Computational requirements for the QM/MM nonadiabatic dynamics simulations.

Molecular dynamics time step	0.05 fs
Real time per dynamics time step (using 3 nodes/96 PEs)	5 s
Length of each trajectory	5000 fs
Number of time steps per trajectory	100,000
Node hours per trajectory (using 3 nodes/96 PEs)	416.7
Number of trajectories	288
Total number of node hours	120,010

D. Estimated Blue Waters Usage Schedule

Both projects will use the allocation evenly throughout the year: Q1: 25%, Q2, 25%, Q3, 25%, Q4, 25%.