Knight, C.; Tse, S.; Savage, J.; Peng, Y.; Voth, G. A.

## Multiscale Problems in Condensed Phase Systems

Many processes important to chemistry, materials science, and biology cannot be described without considering electronic and nuclear-level dynamics and their coupling to slower, cooperative motions of the system. These inherently multiscale problems require computationally efficient and accurate methods to converge statistical properties. The simulation of these reactive processes is not possible with typical molecular mechanics forcefields, which assume a fixed bonding topology throughout the simulation preventing bond formation and cleavage. Condensed phase simulations with electronic structure methods, which naturally account for variable bond topologies, are currently possible, but the computational cost to sample the required time and length scales with sufficient statistics is too prohibitive. It would thus be highly advantageous to have a general methodology for the development of empirical reactive force fields derived from condensed phase electronic structure calculations. These reactive models would reproduce the thermodynamic and dynamic properties of the original *ab initio* calculations, but at a fraction of the computational cost enabling one to extend the time and length scales sampled. This work will discuss recent developments in theories, methodologies and computer algorithms to construct multiscale reactive models derived from condensed phase *ab initio* molecular dynamics simulations (AIMD) or quantum mechanics/molecular mechanics (QM/MM) hybrid methods.
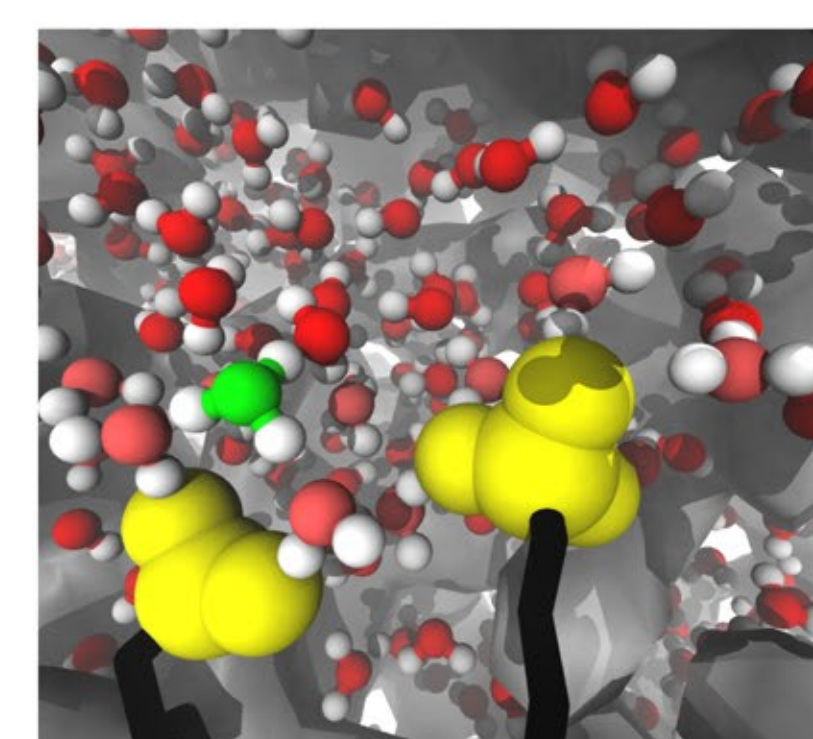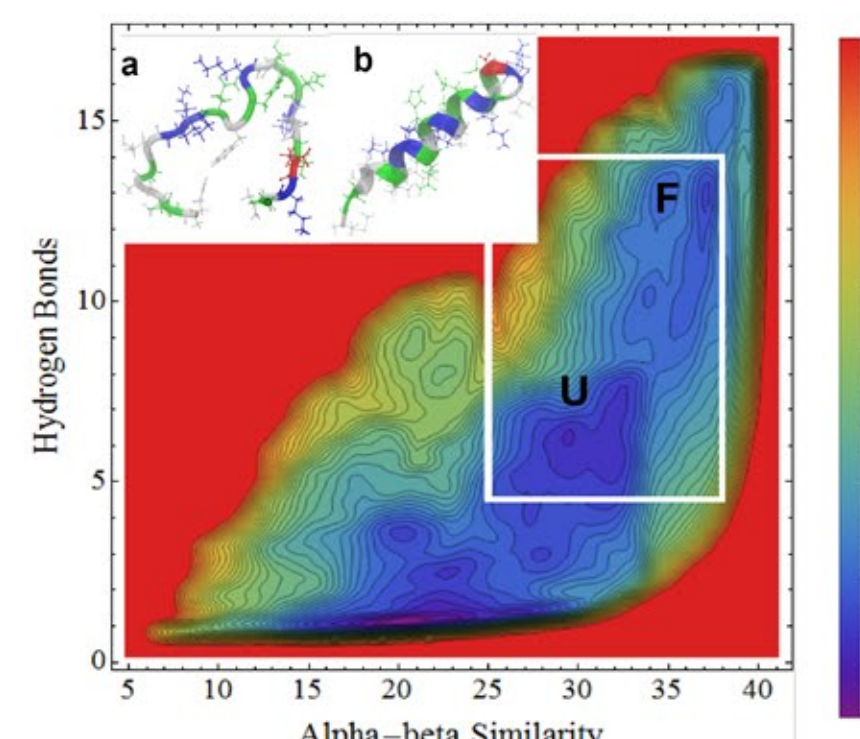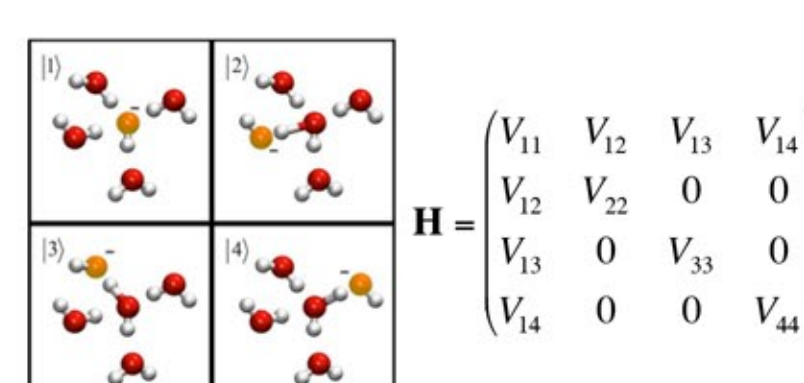

Polymer Membrane
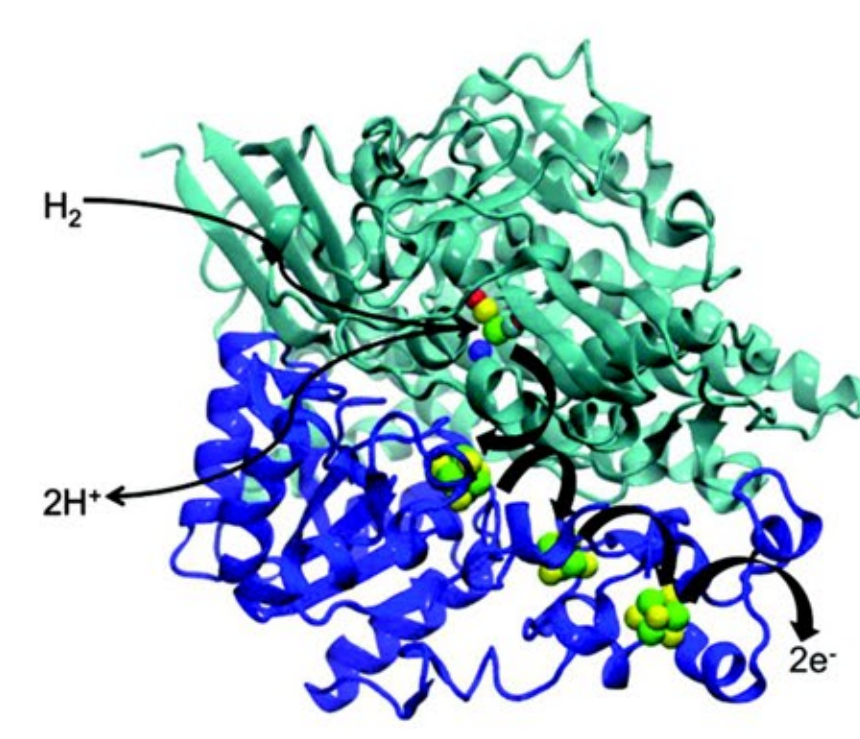
Channels and Enzymes
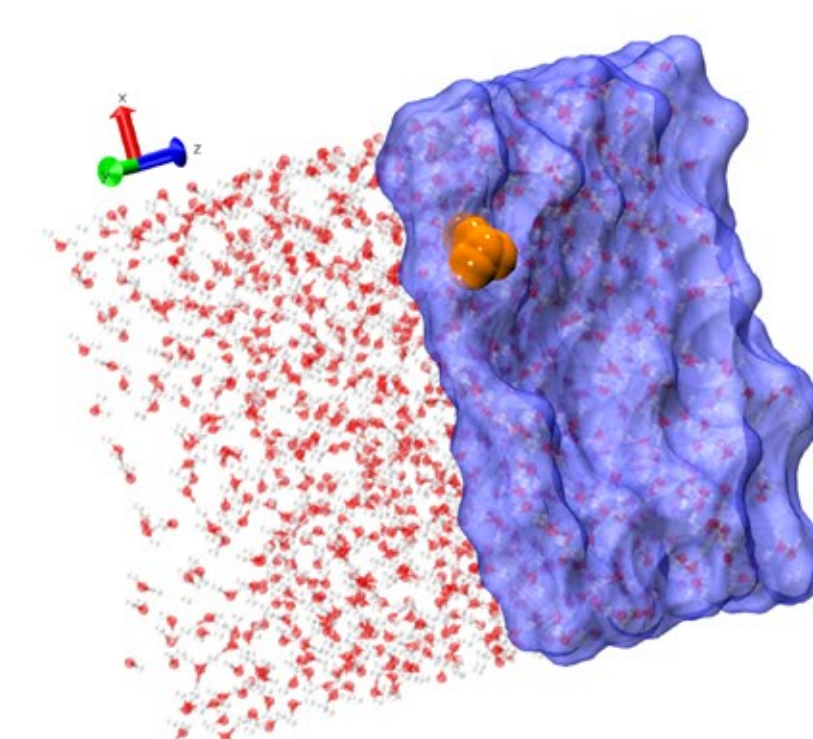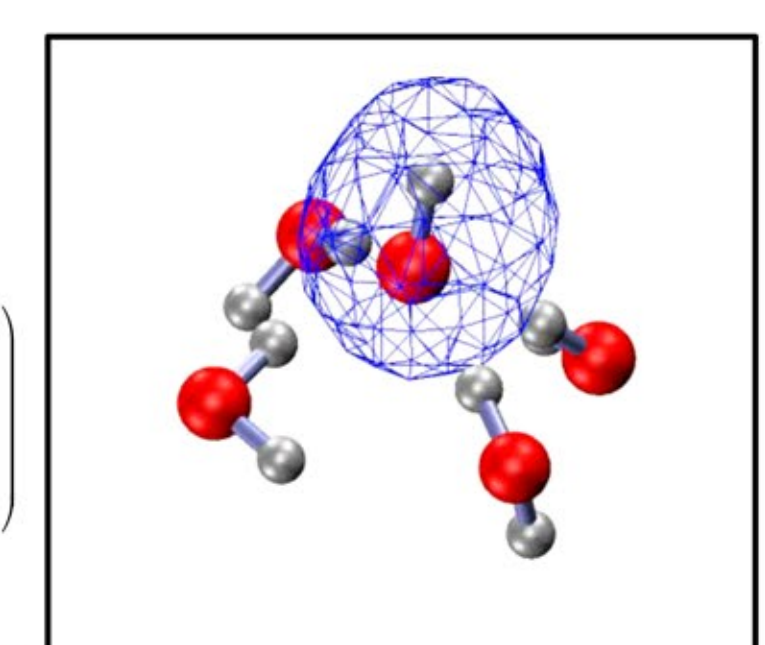
Liquid-Solid Interfaces

Folding and Unfolding

Theory and Methodology

- A linear combination of bonding topologies can be used to describe the *dynamic* bond topology of a reactive complex.

$$H = \begin{pmatrix} V_{11} & V_{12} & V_{13} & V_{14} \\ V_{12} & V_{22} & 0 & 0 \\ V_{13} & 0 & V_{33} & 0 \\ V_{14} & 0 & 0 & V_{44} \end{pmatrix}$$
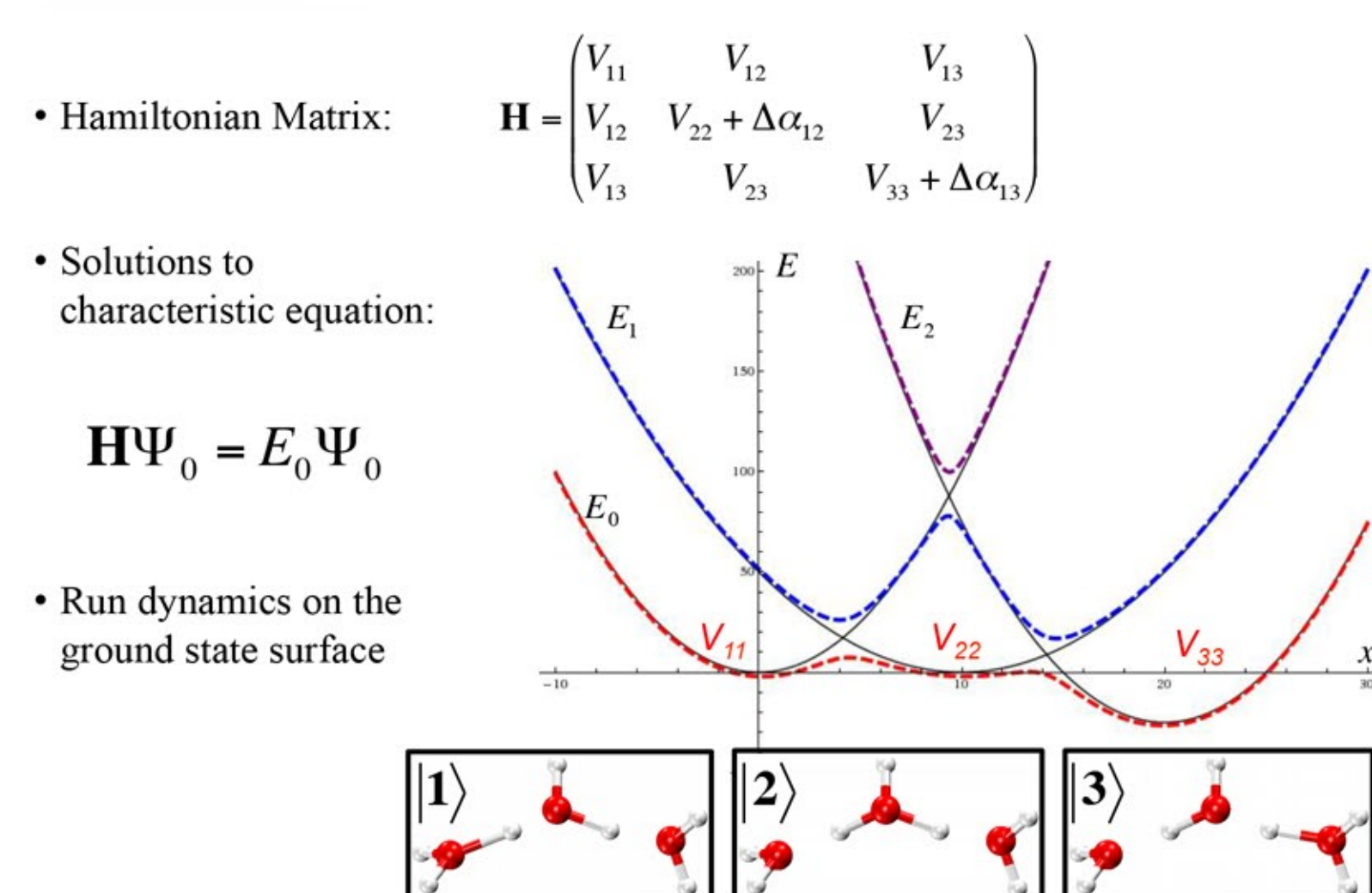
- Diabatic States: $V_{ii} = V_{MM} + V_{REP}$
- Couplings: $V_{ij} = (V_{CONST} + V_{EX})A_{GED}$

All interactions are parameterized by fitting to QM or AIMD data.

Knight and Voth, Acc. Chem. Res., 45, 101 (2012)
Knight, Lindberg, and Voth, J. Chem. Phys., 137, 22A525 (2012)
Knight, Maupin, Izvekov, and Voth, J. Chem. Theory Comput. 6, 3223 (2010)

A linear combination of multiple bonding topologies (diabatic states) is used to describe the reactive system, and to construct a quantum-like Hamiltonian matrix. The diabatic states are determined "on the fly" over the course of a simulations and thus this method is able to account for the dynamic chemical bond rearrangements.
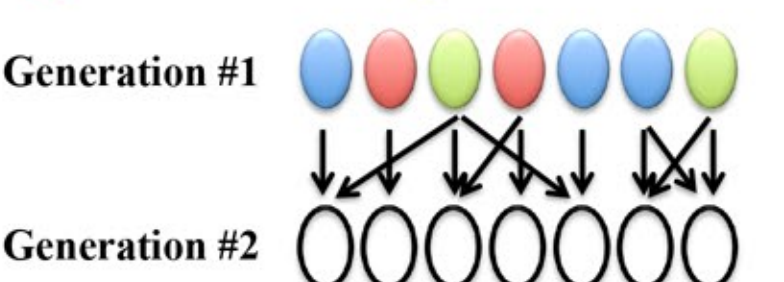
- Hamiltonian Matrix:

$$H = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{12} & V_{22} + \Delta\alpha_{12} & V_{23} \\ V_{13} & V_{23} & V_{33} + \Delta\alpha_{13} \end{pmatrix}$$

- Solutions to characteristic equation:

$$\mathbf{H}\Psi_0 = E_0\Psi_0$$

- Run dynamics on the ground state surface

The off-diagonal elements describe the coupling between two diabatic states (donor and acceptor) and provide a mechanism for transitions, allowing for the donor and acceptor to switch identities. When properly parameterized with condensed phase *ab initio* calculations, the reactive models can be thought of as an extension of the original electronic structure calculation in which the electronic density have been integrated away yielding a **coarse-grained, effective** model for the nuclei.

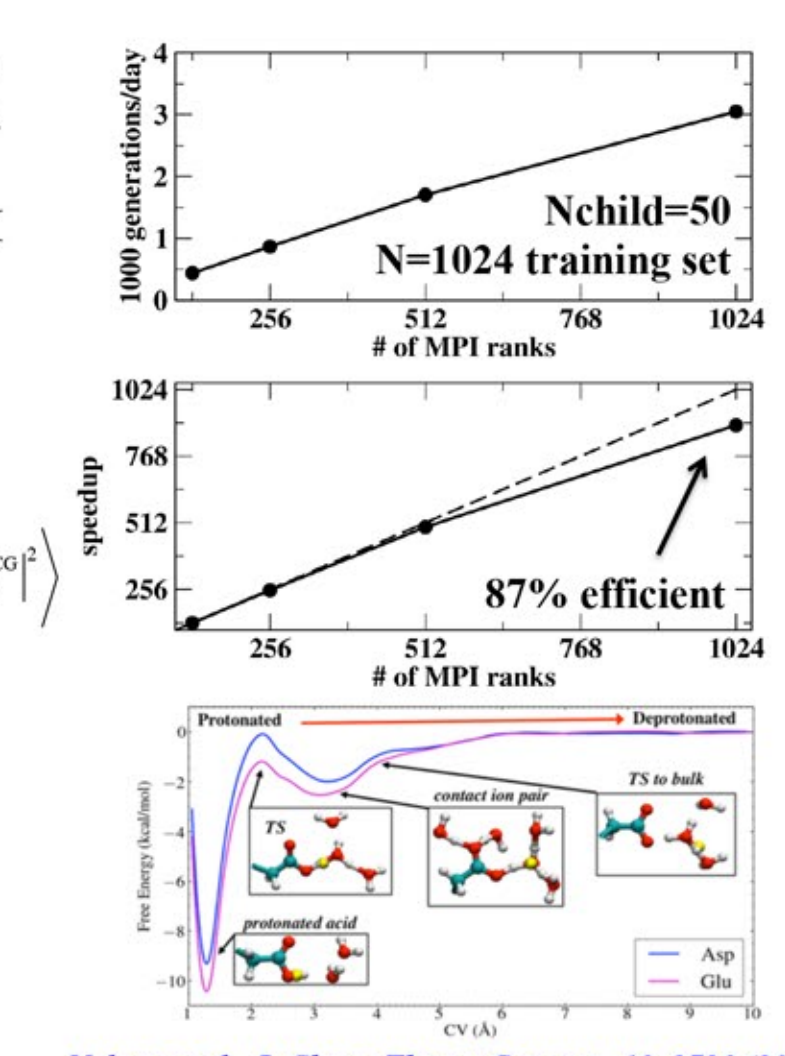- The FitEVB code is tailored to fitting multistate models: Genetic Algorithm.
- Configurations distributed across MPI ranks.
- Each MPI rank evaluates all children for owned configurations.

Normal Child | Mutation | Elite Child

Generation #1

Generation #2

$$\chi^2 = \frac{1}{3N}\sum_{j=1}^{N}\left(\mathbf{F}_j^{AIMD} - \mathbf{F}_j^{CG}\right)^2$$

Nchild=50
N=1024 training set

87% efficient

Nelson et al., J. Chem. Theory Comput., 10, 2720 (2014).

The goal of parameterization is to minimize the difference residue between the forces calculated from the *ab initio* methods and reactive molecular dynamics. It is advantageous to have a flexible and systematic algorithm for parameterizing the reactive models based on AIMD or QM/MM results. To efficiently perform the modeling procedure through fitting a large number of AIMD or QM/MM configurations, A highly parallelized fitting code using genetic algorithm (FitEVB) has been developed.

RAPTOR (Rapid Approach for Proton Transport and Other Reactions) is an *in-house* program developed as a plug-in package to the LAMMPS simulation software and for the multiscale reactive molecular dynamics simulations. This section presents different approaches that have been developed to improve the computational efficiency and parallel scalability of this code in recent years.
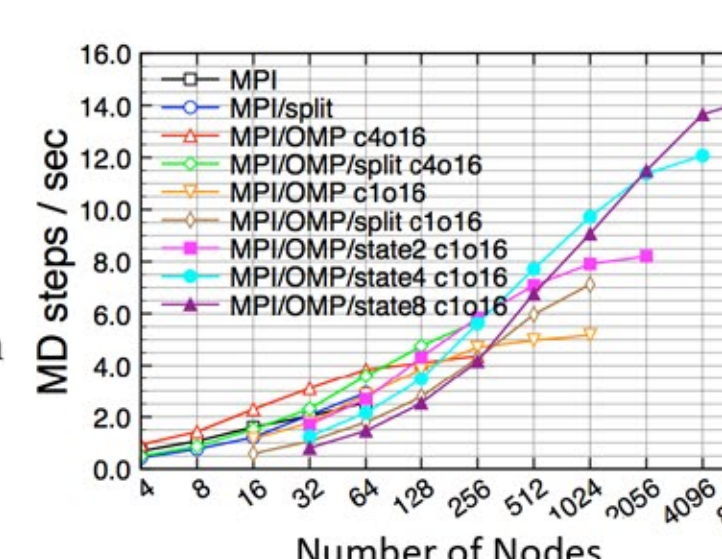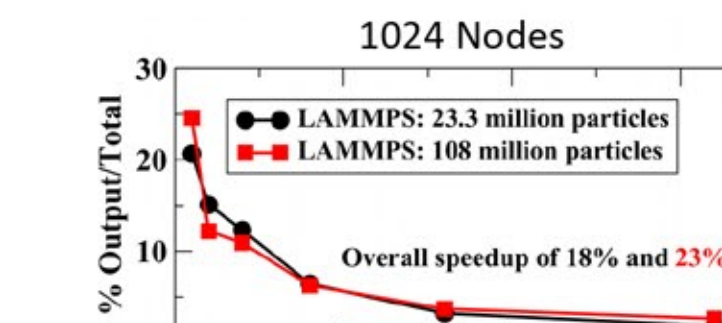
RAPTOR optimization in LAMMPS

- Multithreading with OpenMP
- Parallel I/O
- State-decomposition parallelization

$$H = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{12} & V_{22} & 0 \\ V_{13} & 0 & V_{33} \\ V_{14} & 0 & 0 \end{pmatrix}$$

- Complex-decomposition parallelization

$$H_1 = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{12} & V_{22} & 0 \\ V_{13} & 0 & V_{33} \end{pmatrix} \cdots H_n = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{12} & V_{22} & 0 \\ V_{13} & 0 & V_{33} \end{pmatrix}$$

Adrian Lange, Chris Knight, Yuxing Peng

The most straightforward optimization for parallel efficiency is using the **OpenMP/MPI hybrid scheme**. Besides, the multi-layer decomposition scheme has been developed to split the the computational tasks (states, complex, etc.) at different levels.
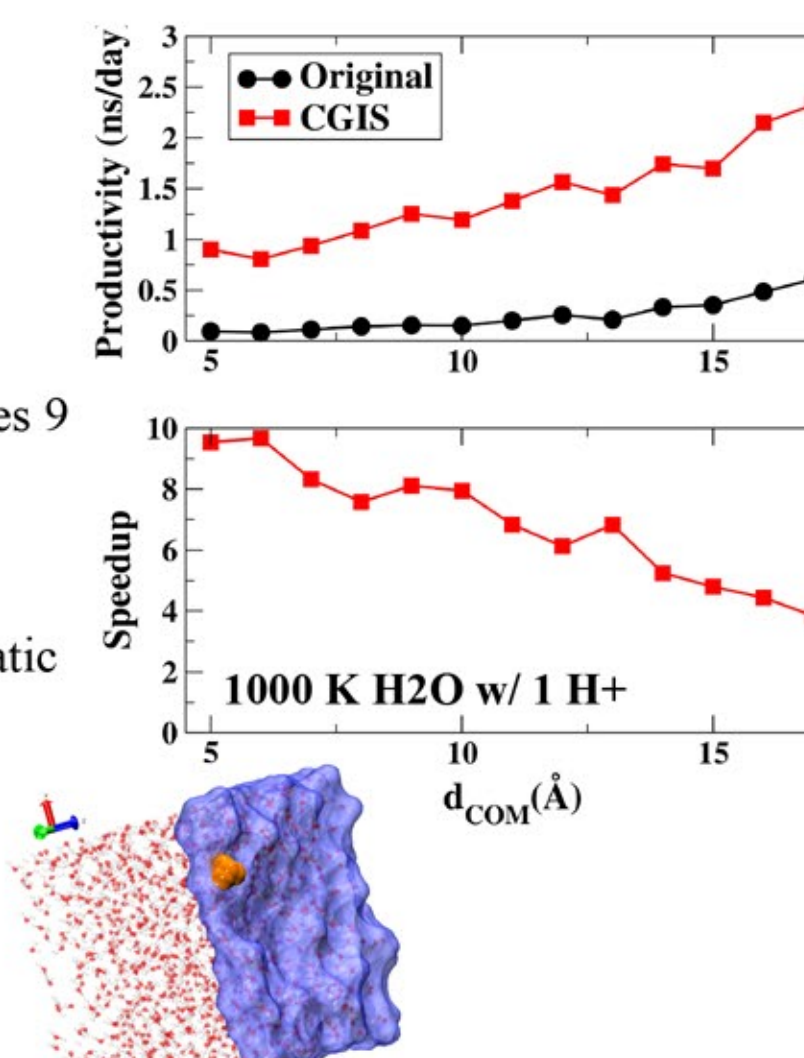
- Is it necessary to use a proper k-space method (PPPM) for off-diagonal electrostatics?

$$E_{A-B} = E_{AB} - E_A - E_B$$

- Off-diagonal interaction energy requires 9 FFT operations.
- Damped, shifted force approximation derived from force-matching electrostatic interactions.
- Hamiltonian dimension decreases as proton approaches interface.

T. Yamashita, Y. Peng, C. Knight, and G.A. Voth, J. Chem. Theory Comput. 8, 4863 (2012)

In MS-RMD simulations, the calculations of the k-space solver (PPPM) for the **long-range electrostatic interactions** takes up more than 70% of the total computational cost. Improving the algorithm for this task can significantly speed-up simulations.

- In many resonance states, the underlying MM model results in negligible weight in ground state wavefunction.
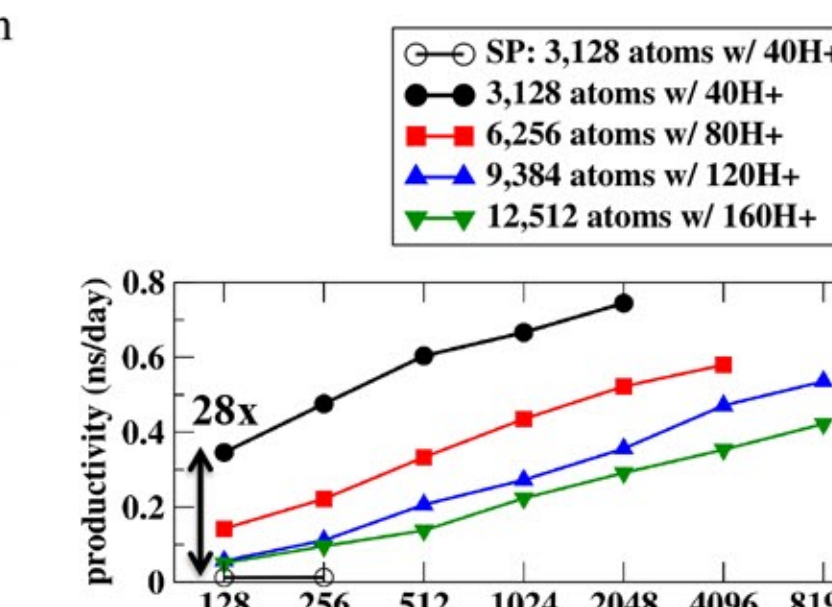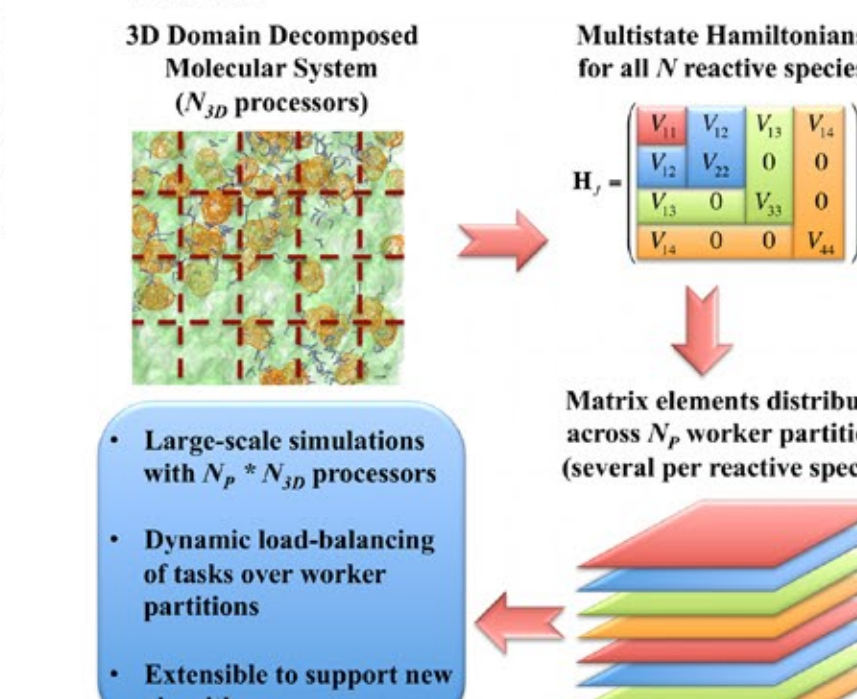- Use local, short-ranged Hamiltonian to efficiently screen unimportant states.

Short Nonbonded Distance

Extended Bond
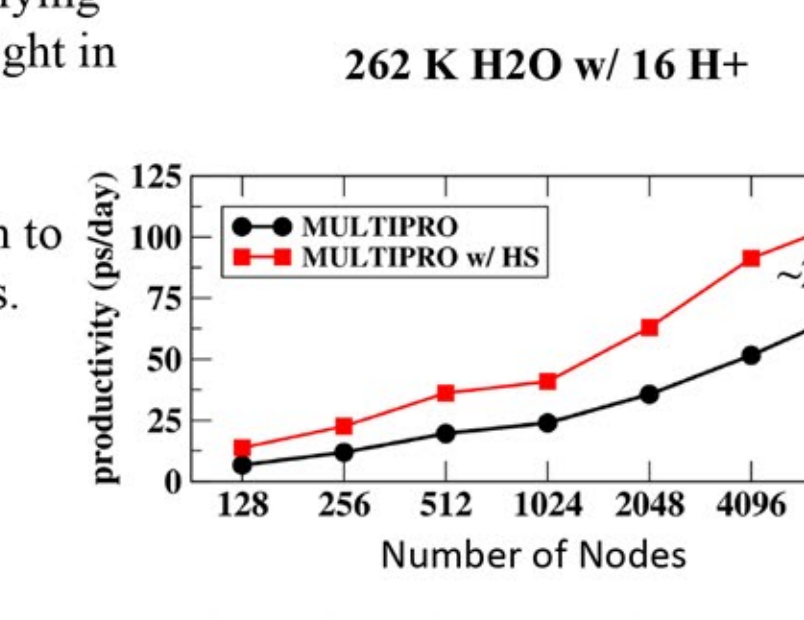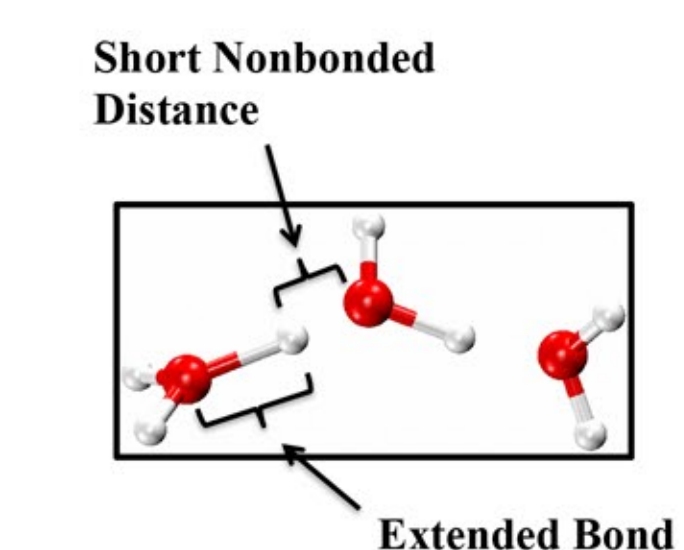
$$H_j = \begin{pmatrix} V_{11} & V_{12} & V_{13} & 0 \\ V_{12} & V_{22} & 0 & 0 \\ V_{13} & 0 & V_{33} & 0 \\ 0 & 0 & 0 & V_{44} \end{pmatrix}$$

The computational cost for each step in the MS-RMD simulations is linear to the number of the diabatic states that is searched. Many of the states have negligible weights to the multistate system. Therefore, a **pre-screening approach** is developed to remove such unimportant states to save the computation time.

- Considerable speedups achieved (even on laptop and local workstations).
- 28x speedup for smallest system compared to Single Partition on 128 nodes.

3D Domain Decomposed Molecular System (N_p processors)

Multistate Hamiltonians for all N reactive species

- Large-scale simulations with $N_p * N_W$ processors
- Dynamic load-balancing of tasks over worker partitions
- Extensible to support new algorithms

- Worker partitions will be assigned as real or k-space partitions for batches of matrix elements.
- Developing new algorithms to accelerate SCF convergence and multiple-time-step methods.

A significant speedup has been achieved through the **state-decomposition scheme.** The whole processor set is firstly divided into sub-workers, each of them conduct the calculations for a sub group of states. Each worker can independently adopt MP-scheme and domain-decomposition for further parallelism.

Applications

Hydroxide solvation and transport in anionic exchange membranes
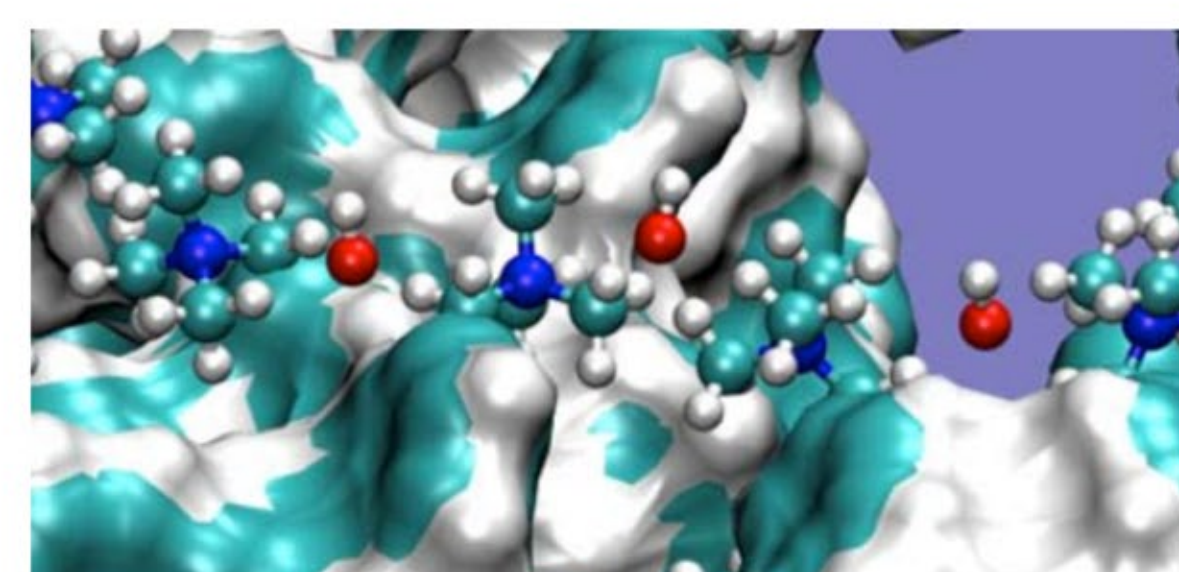
Despite the advantages, anionic fuel cells have not replaced protonic fuel cells primarily because of low ionic conductivity, poor chemical stability and carbonate formation. This work attempts to understand hydroxide solvation and transport in anionic fuel cells by atomic level modeling and suggest improved design strategies from that microscopic information.

Persistent Subdiffusive Proton Transport in PFSA Membranes

We show that PT is indeed subdiffusive for several hundred picoseconds for all hydration levels examined, and the it is highly dependent on water concentration. We also investigate the caging of the excess proton using a recently developed technique and show that the excess proton exhibits caging effects up to at least 1 ns in PFSA systems.
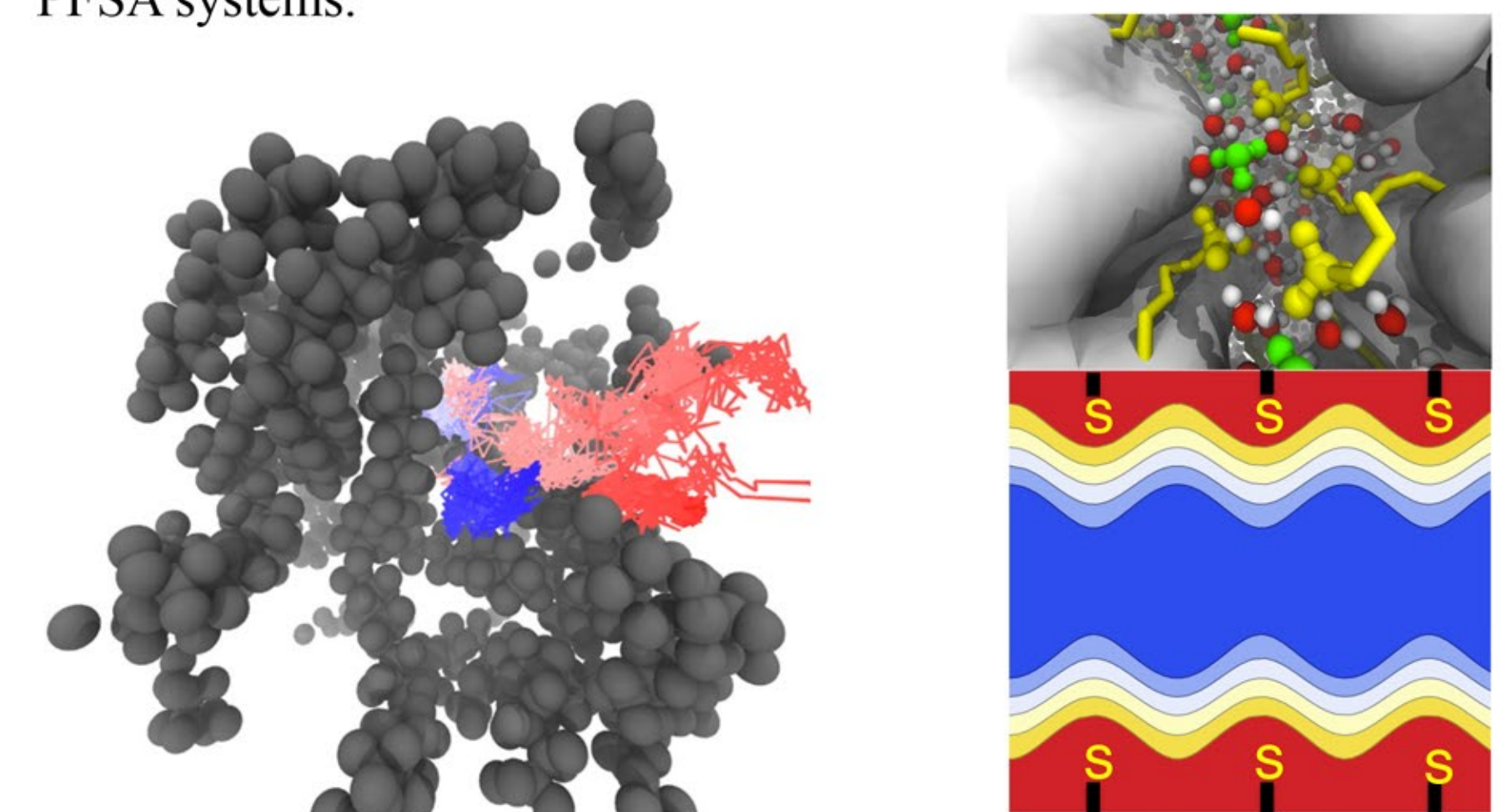
Effects of Polymer Morphology on Proton Solvation and Transport in Proton-Exchange Membranes

The effects of polymer morphology on proton solvation and transport in hydrated Nafion are investigated by using a novel reactive molecular dynamics approach. the interaction between protons and the sulfonate groups is shown to be the key factor in determining PT behavior.
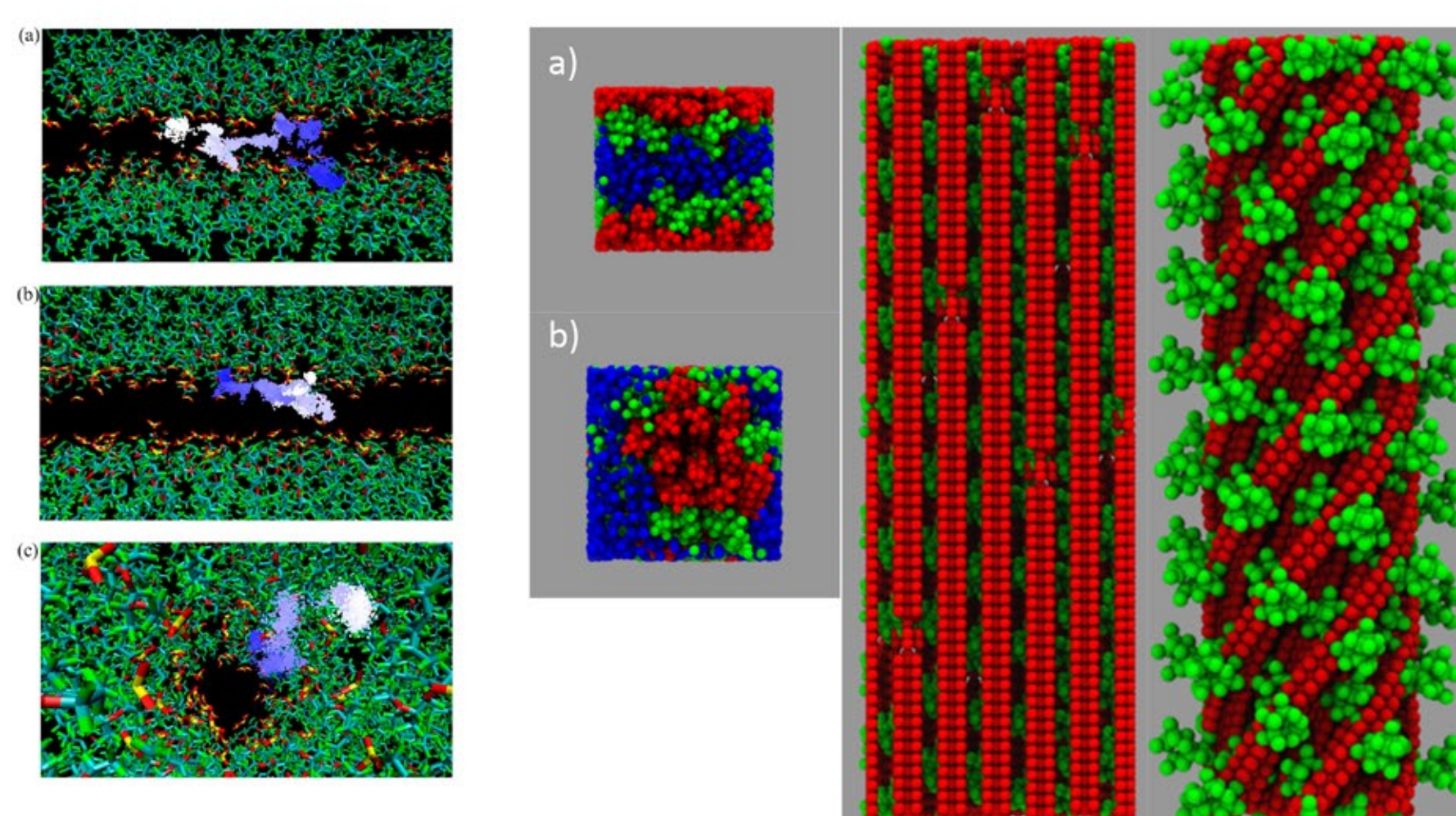
Proton Transport Pathways in [NiFe]-Hydrogenase

A key step in the hydrogenase catalytic cycle and the focus of this work is proton transport (PT) to and from the active site. The PT mechanism of the enzyme is studied using reactive molecular dynamics simulations of the full protein and the excess proton transfers via the multistate empirical valence bond (MS-EVB) method. Pathways connecting the bulk and the active site are located that suggest possible participation by several protonatable residues. PT free energy surfaces are calculated to differentiate the pathways.
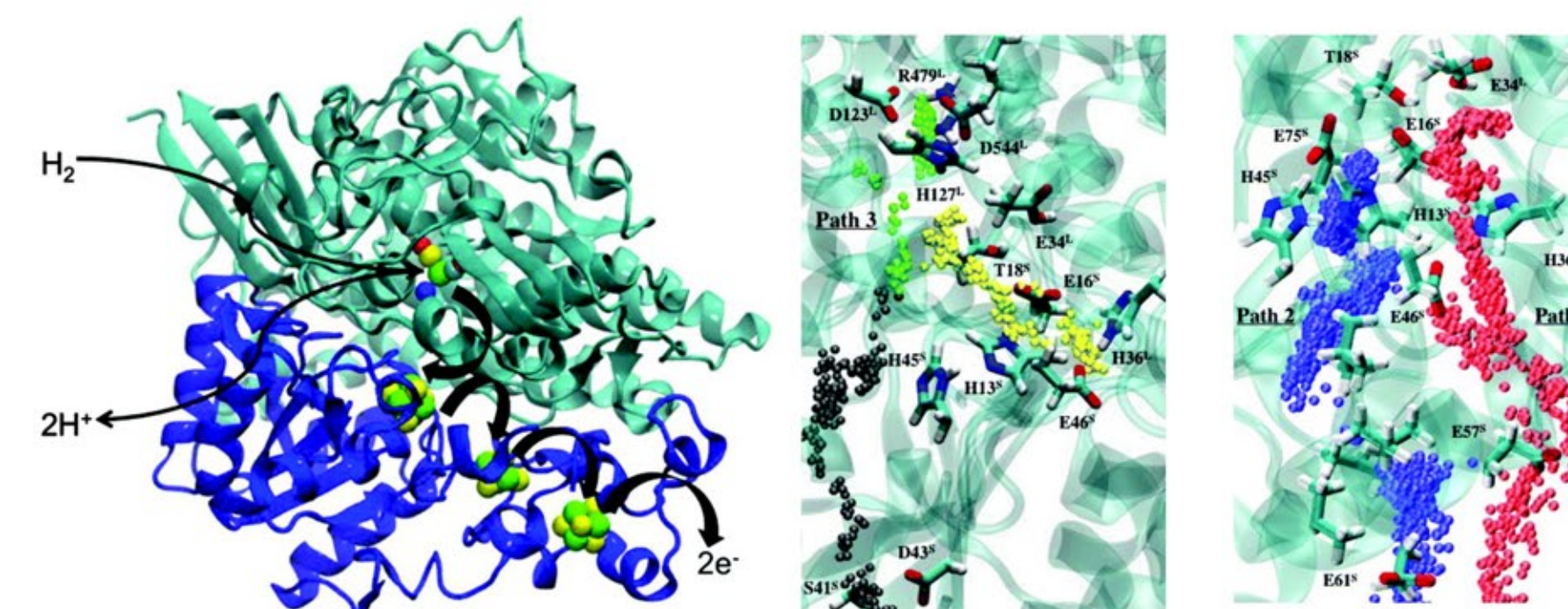
Proton Induced Wetting in Hydrophobic Spaces
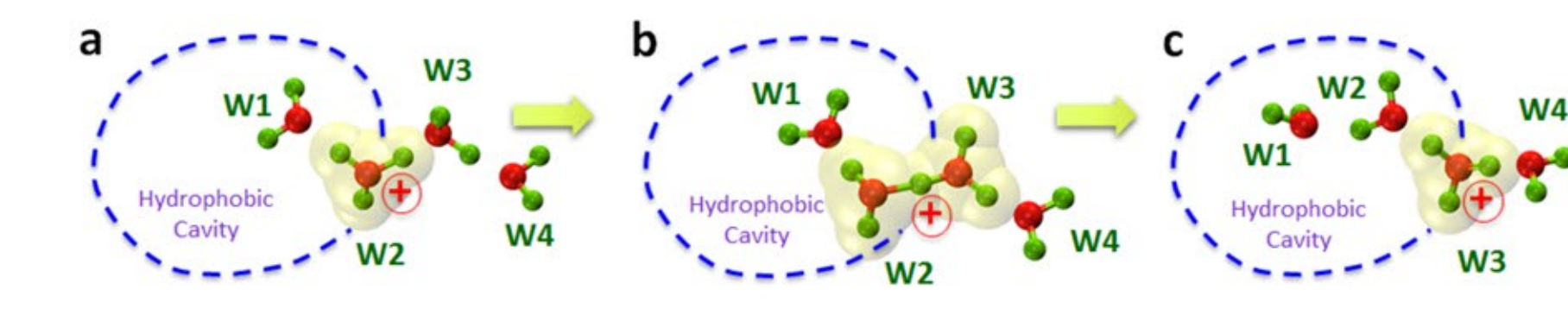
For many years, *Grotthuss* proton transport through hydrogen bonded water networks has been the focus of theoretical and experimental discussions. However, is there a related process in which water can be transported through a proton charge defect in order to "wet" the "path ahead" for the charge defect to then migrate? This process has been recently observed in our computer simulations of proton transport through a hydrophobic carbon nanotube.