

Quantifying Uncertainties in Solvation Procedures for Modeling Aqueous Phase Reaction Mechanisms

Alex M. Maldonado,[†] Satoshi Hagiwara,[‡] Tae Hoon Choi,[†] Frank Eckert,[¶]
Kathleen Schwarz,[§] Ravishankar Sundararaman,^{||} Minoru Otani,[‡] and John A.
Keith^{*,†}

[†]*Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh,
Pennsylvania 15261, United States*

[‡]*National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono,
Tsukuba 305-8568, Japan*

[¶]*Dassault Systèmes Deutschland GmbH, Imbacher Weg 46, 51379 Leverkusen, Germany*

[§]*Material Measurement Laboratory, National Institute of Standards and Technology,
Gaithersburg, Maryland 20899, United States*

^{||}*Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy,
New York 12180, United States*

E-mail: jakeith@pitt.edu

Abstract

Computational quantum chemistry provides fundamental chemical and physical insights into solvated reaction mechanisms across many areas of chemistry, especially in homogeneous and heterogeneous renewable energy catalysis. Such reactions may depend on explicit interactions with ions and solvent molecules that are nontrivial to characterize. Rigorously modeling explicit solvent effects with molecular dynamics usually brings steep computational costs while the performance of continuum solvent models such as polarizable continuum model (PCM), charge-asymmetric nonlocally determined local-electric (CANDLE), conductor-like screening model for real solvents (COSMO-RS), and effective screening medium method with the reference interaction site model (ESM-RISM) are less well understood for reaction mechanisms. Here, we revisit a fundamental aqueous hydride transfer reaction—carbon dioxide (CO_2) reduction by sodium borohydride (NaBH_4)—as a test case to evaluate how different solvent models perform in aqueous phase charge migrations that would be relevant in general to renewable energy catalysis mechanisms. For this system, quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations almost exactly reproduced energy profiles from QM simulations, and the Na^+ counterion in the QM/MM simulations plays an insignificant role over ensemble averaged trajectories that describe the reaction pathway. However, solvent models used on static calculations gave much more variability in data depending on whether the system was modeled using explicit solvent shells and/or the counterion. We pinpoint this variability due to unphysical descriptions of charge-separated states in the gas phase (i.e., self-interaction errors), and we show that using more accurate hybrid functionals and/or explicit solvent shells lessens these errors. This work closes with recommended procedures for treating solvation in future computational efforts in studying renewable energy catalysis mechanisms.

Introduction

Energy consumption across the world has increased on average by 1.6 % every year since 2008, and this annual increase is approximately equal to what the United Kingdom uses each year.¹ Sustainability concerns thus drive research in renewable energy catalysis based on enzymes² or (photo-)electrochemistry,^{3,4} and reaction mechanisms for these processes often involve the intricate participation of solvent molecules and ions that are nontrivial to model.⁵

Computational quantum chemistry (QC) allows reliable predictions of these reaction mechanisms, but rigorously treating the dynamic nature of electrolytes with quantum mechanics (QM) based molecular dynamics⁶ or quantum mechanics/molecular mechanics (QM/MM) sampling procedures⁷ can bring high computational costs when investigating viable chemical reaction pathways in solvents. Alternatively, continuum solvent models (CSMs) such as the conductor-like screening model for real solvents (COSMO-RS),⁸ conductor-like polarizable continuum model (CPCM),⁹ effective screening medium method with the reference interaction site model (ESM-RISM),¹⁰ and charge-asymmetric nonlocally determined local-electric (CANDLE)¹¹ have been developed to allow computational investigations with far greater efficiency. The reliability of these methods, even when used with small numbers of explicit solvent molecules, is still a topic of ongoing research.¹²⁻¹⁵ See Ref. 16 for a recent review on applying continuum and explicit approaches to describe solvation and field effects in first-principles electrochemistry. However, no report has yet shown a systematic and comparative assessment of all these methods on *the same reaction mechanism*.

Reduction of carbon dioxide (CO₂) into renewable fuels and chemicals remains a key process in renewable energy catalysis.^{17,18} Grice et al. experimentally demonstrated that sodium borohydride is sufficiently hydritic to reduce CO₂ even in aqueous solutions.¹⁹ We previously used this system to computationally study model sensitivities in this fundamental charge migration process using static^{20,21} and dynamic²² calculation schemes. For static calculations, Groenenboom and Keith obtained explicit solvating environments with high

temperature Born-Oppenheimer molecular dynamics (BOMD) simulations with 70 water molecules to analyze borohydride hydrolysis mechanisms without any *a priori* assumptions.²¹ These automatically predicted reaction events were followed with generalized solid-state nudged elastic band (g-SSNEB) calculations to automatically predict the reaction’s minimum energy pathway.²³ The g-SSNEB images were then be decomposed into subsystems of atoms that delineated the first explicit solvent shell with or without the local counterion (see Figure 1).

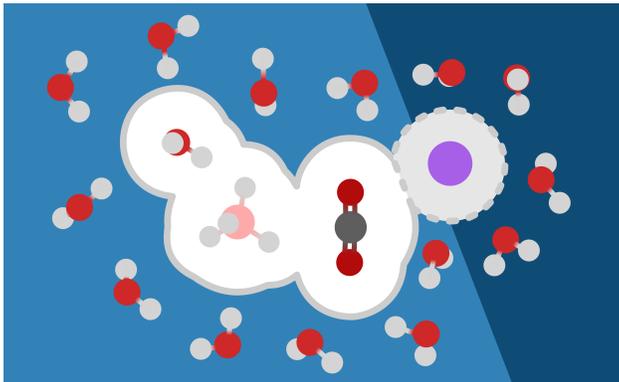


Figure 1: Illustrations of the four g-SSNEB subsystems studied in this and prior work.²¹ The smallest subsystem contains only the molecules directly participating in the reaction: BH_4^- , CO_2 , and H_2O (surrounded by a solid gray line). This subsystem can also include a Na^+ counterion shown with a dashed gray circle. The remaining two subsystems involve the first explicit solvent shell with the reacting molecules and the corresponding solvent molecules coordinated to the Na^+ counterion.

Without the first explicit solvent shell or the counterion present, solvation model based on density (SMD)²⁴ and conductor-like screening model (COSMO)²⁵ could not accurately predict local solvent contributions to the reaction energies. Computationally intensive potential of mean force (PMF) calculations using QM simulations along the same reaction coordinate as the g-SSNEB calculation provided rigorous, dynamic insights into the reaction.²² It was found that free energy profiles were sometimes in good agreement with reaction energies from static calculations as long as one of the endpoints for the g-SSNEB pathway was not trapped in a metastable state.

Several unanswered questions remain that could help provide insight into how to best

model catalytic reaction mechanisms under solvating conditions. For example, how different are costly QM simulations compared to larger, more efficient QM/MM simulations that may or may not include the explicit counterion? Furthermore, semiempirical QM methods (e.g., GFN*n*-xTB²⁶) have recently been designed to efficiently predict equilibrium structural properties involving atoms across the entire periodic table. Rapid QC models would provide tremendous cost advantages for thorough sampling of molecular configurations, but it is also not clear how well they would perform in solvated reaction mechanisms.

CSMs have a variety of mathematical implementations for solute cavities and solute-solvent interactions¹⁶ ranging from electrostatic interactions with a dielectric or conducting cavity (including SMD, CANDLE, COSMO and COSMO-RS) to ones that incorporate solvent structure (such as ESM-RISM). Some CSMs are more complex (e.g., CANDLE, COSMO-RS, and ESM-RISM) than others in terms of formulations and/or applications. Thus, open questions include: how differently do these approaches perform for a single reaction mechanism study, and how do these perform in a fundamental hydride transfer reaction mechanism?

Finally, we wanted to address the importance of explicit solvent methods in reaction mechanism studies. Can a reaction be reliably modeled using automatable procedures using a system described by a few explicit solvent molecules and little to no *a priori* expectation for what intermediate and product states will form? These scientific questions guided our collaborative effort that spanned multiple research groups. To help advance understanding of these calculations, the supporting information (SI) provides extensive documentation and output files to allow the reproduction of this work for further study and training purposes.

Methods

QM/MM molecular dynamics simulations

QM/MM molecular dynamics (QM/MM MD) simulations were performed in GAMESS²⁷ and umbrella sampling was used to sample along the aqueous phase reaction.¹ Simulations were performed with the spherical boundary condition surrounded by 285 to 290 water molecules to form a water sphere with a 11 Å radius. A harmonic restraint potential with a force constant of 2.0 kcal/mol/Å² (0.087 eV/Å²) was applied to keep the volume of sphere constant. The QM and MM regions were described with ω B97X-D/6-31G(d)^{28,29} and TIP5P³⁰ water model, respectively. The simulation temperature of 300 K was kept constant by the Nosé–Hoover thermostat in the NVT ensemble. A 20 ps equilibration was performed on each window with a time step of 1 fs; production runs of over 50 ps were continued from the final equilibration structures. The PMF from the umbrella samplings were obtained using the weighted histogram analysis method (WHAM).³¹

Static cluster calculations

Explicitly solvated structures were obtained from our previous work²¹ where QM BOMD simulations under periodic boundary conditions using the PBE exchange correlation functional³² and the projector augmented wave (PAW) method^{33,34} at high temperatures provided reactant and product structures that were later used for g-SSNEB optimizations at 0 K. Note that structures for all reaction steps in the mechanism are modeled using the same number of atoms. This causes contributions arising from ideal gas, rigid rotor, harmonic oscillator (IGRRHO) approximations to generally cancel out on a relative scale.

Gas-phase electronic energies were calculated using using ORCA 4.2.0.^{35,36} The cclib package³⁷ was used to parse data from calculations when possible. We compared differ-

¹Note: Certain commercial materials are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified is necessarily the best available for the purpose.

ent exchange-correlation functionals including the PBE and BP86^{38,39} generalized gradient approximation (GGA) functionals, the B3LYP hybrid functional,^{38,40-42} the ω B97X range-separated hybrid functional,²⁸ and the domain-based local pair natural orbital approach, DLPNO-CCSD(T) calculations.⁴³ For Kohn-Sham density functional theory (DFT) calculations we compared the performance of the Ahlrichs basis sets,⁴⁴ mostly using the relatively small def2-SVP and larger def2-TZVP basis sets, while the def2-QZVPP basis set was used for DLPNO-CCSD(T). We also compared the above calculations to high quality complete basis set (CBS) extrapolated CCSD(T)⁴⁵ as implemented in ORCA.^{46,47} Two-point extrapolation using cc-pVTZ and cc-pVQZ basis sets^{48,49} were used with an alpha and beta of 5.460 and 3.050, respectively.^{50,51}

Since DFT explicitly neglects dispersion interactions, atom-pairwise dispersion corrections using the Becke-Johnson damping scheme (D3BJ)^{52,53} were used. DFT calculations also used the ‘tight’ self-consistent field (SCF) convergence criteria, and the Lebedev-302 angular grid with 4.67 general integration accuracy (Grid4) as defined in ORCA. The RI-J approximation with def2/J basis set⁵⁴ were included for PBE and BP86 calculations (default in ORCA); this approximation was confirmed to have a negligible effect on the accuracy when modeling g-SSNEB subsystems. GFN1-xTB⁵⁵ and GFN2-xTB⁵⁶ gas-phase electronic energies were computed with xtb package.⁵⁷

Solvation energy calculations

Solvation energies from CSMs were calculated as the difference between liquid- and gas-phase electronic energies at the same level of theory. All CSMs use default parameters for water unless specified otherwise. Energies from CPCM⁹ and SMD²⁴ were calculated using ORCA. We also compared these results to solvation energies obtained using ω B97X-D/def2-TZVP calculations with the polarizable continuum model (PCM)⁵⁸ as implemented in Gaussian 16 Rev. C.01.⁵⁹ CANDLE solvent model¹¹ calculations with a 1.0 mol/L electrolyte were from the JDFTx 1.5.0 code.⁶⁰ CANDLE determines the solvation cavity from a convolution

of the solute and solvent electron densities,⁶¹ and it accounts for the variation of solvation cavity with solute charge. The combination of nonlocal cavity determination with charge asymmetry enables accurate solvation of neutral solutes, cations and anions within a single parameterization.¹¹

The BIOVIA COSMOtherm 2020 package⁶² was used to perform COSMO-RS computations using approach-specific BP/TZVP/COSMO and BP/TZVPD/FINE levels⁶³ for water, 0.1 mol/L, and 1.0 mol/L NaBH₄. The QC COSMO calculations⁶⁴ were performed with the Turbomole code.⁶⁵ Single-point energy calculations used BP86 with the def-TZVP and def2-TZVPD basis sets,⁶⁶ respectively. For a justification for the choice of functional and basis set used for the COSMO approach, see Ref. 67. To calculate the molecular surfaces, the default radii from COSMO,⁶⁸ as implemented in the Turbomole program package, were used. Further details of the DFT/COSMO calculations with Turbomole are given in Ref. 64. The screening charge surfaces of the solvent molecules on the BP/TZVP/COSMO and BP/TZVPD/FINE levels were taken from BIOVIA COSMObase 2020,⁶⁹ a database of pre-computed COSMO surfaces.

ESM-RISM calculations were carried out using plane-wave basis sets within the ultra-soft pseudopotential framework^{70,71} implemented in Quantum ESPRESSO code (PWSCF 6.1).^{72,73} The cut-off energies for wave functions and augmented charges were 40 Ry (544.23 eV) and 320 Ry (4353.82 eV), respectively. Only the gamma point was used for k -point sampling. A box size of $30 \times 30 \times 30 \text{ \AA}^3$ was used for the unit cells. The spin-unpolarized PBE exchange-correlation functional was used. Aqueous NaBH₄ solution at a temperature of 298 K was represented by the RISM. Concentration of H₂O solvent was 1.0 cm³/g (55.6 mol/L), and 0.1 mol/L or 1.0 mol/L was used as the concentration of NaBH₄ salt. In the ESM-RISM calculation, both sides of z -direction of unit cells were expanded by 31.75 Å. The relationship between DFT and expanded cells are found in Ref. 74. The cut-off energy for the reciprocal representation of RISM equation was 160 Ry (2176.92 eV). To solve the RISM equation, the Kovalenko and Hirata type closure function was used,⁷⁵ and the solvation free energy

was calculated using the Gaussian fluctuation method. 1.0×10^{-6} Ry (1.36×10^{-5} eV) was used for convergence criteria of correlation functions in the RISM equation. Lennard-Jones (LJ) type classical force field was used, and the LJ parameters and charges used for RISM calculations were selected from literature^{74,76-79} as listed in Table S1. LJ parameters for the heterotype atomic pairs were determined by simple-combination rules.

Growing string method calculations

Single-ended⁸⁰ and double-ended⁸¹ GSM⁸² calculations were performed with the molecularGSM package.⁸³ A gradient convergence tolerance of 1×10^{-4} kcal/mol/Å (4.336×10^{-6} eV/Å), intermediate detection of 2.0 kcal/mol (0.087 eV), and maximum of 30 nodes were used. GSM energy and gradients were from ORCA and employed B3LYP-D3BJ/def2-SVP, tight SCF convergence, and Grid4. All driving coordinates (e.g., breaking one of the B–H bonds and forming a H–C bond) were enumerated over and its products categorized. Initial structures for single-ended GSM calculations were found by using a multistep procedure. First, the lego module in ABCluster 1.5.1^{84,85} was used with ORCA BP86-D3BJ/def2-SVP optimizations (with RI-J approximation) was used to identify between 20 and 25 candidate structures. Structures that appeared visually different were optimized and confirmed to be a stationary point once hessian calculations using the same model chemistry mentioned above resulted in no imaginary frequencies. The structure coinciding with the lowest electronic energy was then selected for single-ended GSM calculations.

Results and Discussion

Comparisons of QM and QM/MM simulations

Our previous study used QM simulations to predict free energy profiles for CO₂ reduction by BH₄⁻ in explicit aqueous solution that contained a Na⁺ counterion.²² We now present data using QM/MM that are much faster, but require some additional technical expertise to run.

Figure 2 shows a comparison of the QM and QM/MM pathway along the same collective variable (g-SSNEB is also provided). The differences in solvated reaction energies for QM

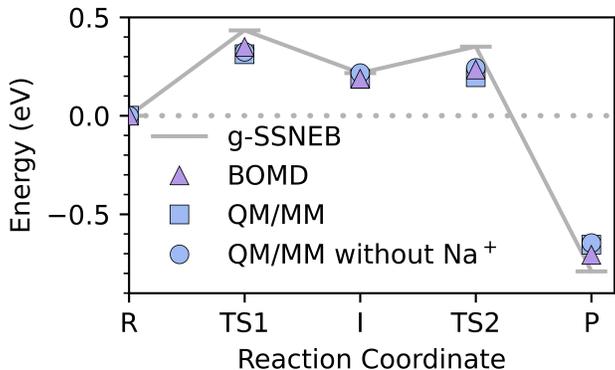


Figure 2: Comparison of QM (PBE/DZVP and GTH pseudopotentials with 70 water molecules)²² and QM/MM MD potential of mean force calculations. The mean absolute error was 0.03 eV with ω B97X-D/6-31G* QM description and 285 TIP5P water molecules. Counterion exclusion insignificantly affects PMF energies.

and QM/MM simulations were quite small which can be attributed to the classical solvent description. For other reaction mechanism studies, the saved cost of QM/MM should be used to explore more reactions, longer simulations (sampling), or higher levels of theory.^{86,87}

Since species involved in a charge migration may be anionic as is the case here, it might be necessary under periodic boundary conditions to include a counterion for charge compensation. Furthermore, in previous work involving static calculations the cation appeared to play a very significant role in stabilizing reaction intermediates.²¹ However, when using QM simulations, the Na⁺ was found to exhibit Brownian motion, indicating that it may not have a significant stabilizing influence on dynamic time scales. The QM/MM approach used here is not run under periodic boundary conditions, so using a neutral charged system is also not necessary. QM/MM simulations without the counterion were found to predict effectively the same energy profile with a mean absolute deviation to QM/MM simulations with the counterion of only 0.02 eV. Thus, we confirm for this system, and likely others, that a counterion should not be needed in dynamic simulations since the role of a counterion is compensated with the stabilizing interactions with solvent molecules over the course of the

simulation trajectories. However, below we will more deeply investigate the impact that the counterion has when using static calculations with and without surrounding explicit solvent shells.

Gas-phase electronic energies

We created a benchmarking data set consisting of g-SSNEB subsystem structures containing the reacting molecules with and without the counterion and single-ended GSM pathways (a total of 16 structures). Errors were calculated relative to CCSD(T)/CBS energies and are presented in Figure 3.

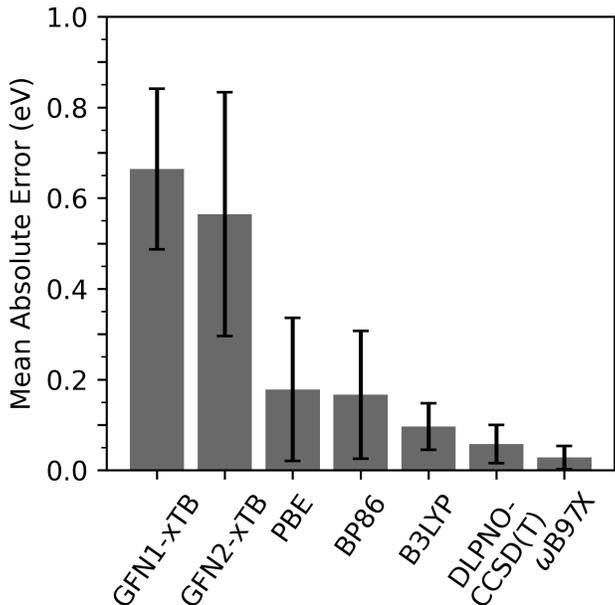


Figure 3: Errors in relative gas-phase electronic energies for reacting molecules (i.e., energy differences of the transition states, intermediates, and products from the reactants) benchmarked against CCSD(T)/CBS calculations. Mean absolute errors of relative energies with standard deviation error bars are presented. All DFT calculations use the def2-TZVP basis set and D3BJ dispersion corrections. DLPNO-CCSD(T) uses def2-QZVPP.

We found that GFN1- (MAE = 0.66 eV) and GFN2-xTB (MAE = 0.57 eV) were considerably inaccurate for these particular boron-containing systems, which is not unexpected since these semiempirical methods were developed for rapid calculations of geometries, vibrational frequencies, and non-covalent interactions.⁵⁵ Thus, we caution against using these

methods outside of their intended use. The GGA PBE and BP86 functionals had the next largest MAEs, 0.18 and 0.17 eV, respectively. The hybrid B3LYP (MAE = 0.10 eV) and the range-separated hybrid ω B97X (MAE = 0.03 eV) performed the most accurately of the tested DFT approaches. In this particular data set, the DLPNO-CCSD(T) results were slightly less accurate than ω B97X-D3BJ results.

This general trend in accuracy of these QC methods agrees with more extensive benchmarking studies that examined thermochemistry, isomerization, non-covalent interaction data sets.⁸⁸ Since ω B97X-D3BJ/def2-TZVP provided the best balance between high accuracy (relative to CCSD(T)/CBS extrapolations) with reasonably lower cost, its results will be used for gas-phase energy contributions in all solvated reaction energies.

Continuum solvent model predictions

As explained above, we used molecular clusters from previous g-SSNEB calculations²¹ to represent static structures within an explicitly solvated system. The reaction pathway calculation found two sequential barrier heights that result in a reactant state (R), the first transition state (TS1), a metastable intermediate state (I), a second transition state (TS2), and a product state (P). Note that this specific pathway reflects just a single configuration of surrounding water molecules and a nearby counterion, and thus it only represents one (of likely very many) possible reaction pathways that would be possible at ambient conditions. **The aim of this section is not to determine the true pathway, but rather to assess how different solvent models reproduce explicit solvent and counterion interactions. Thus, no structures are optimized after removing explicit water molecules to keep the same reactant and counterion configurations throughout all subsystems. While the systems could no longer be in true minima or saddle points, the solvation energy contributions can be directly compared. Modeling the pathway with optimized structures and CSMs is shown later.** Additionally, a harmonic estimation of the tunnelling cross-over temperature for this process⁸⁹ showed that nuclear tunneling is unlikely to play a significant role at the considered temperatures.

Solvated reaction energy profiles using ω B97X-D3BJ/def2-TZVP gas-phase energies and SMD, CPCM, CANDLE, COSMO-RS Fine, and ESM-RISM contributions are shown here. Data using PCM and LinearPCM methods (the latter calculated using JDFTx) were found to be generally similar and are included in the SI. We note that the ionic strength used in the CSM did not play a large role in this reaction pathway.

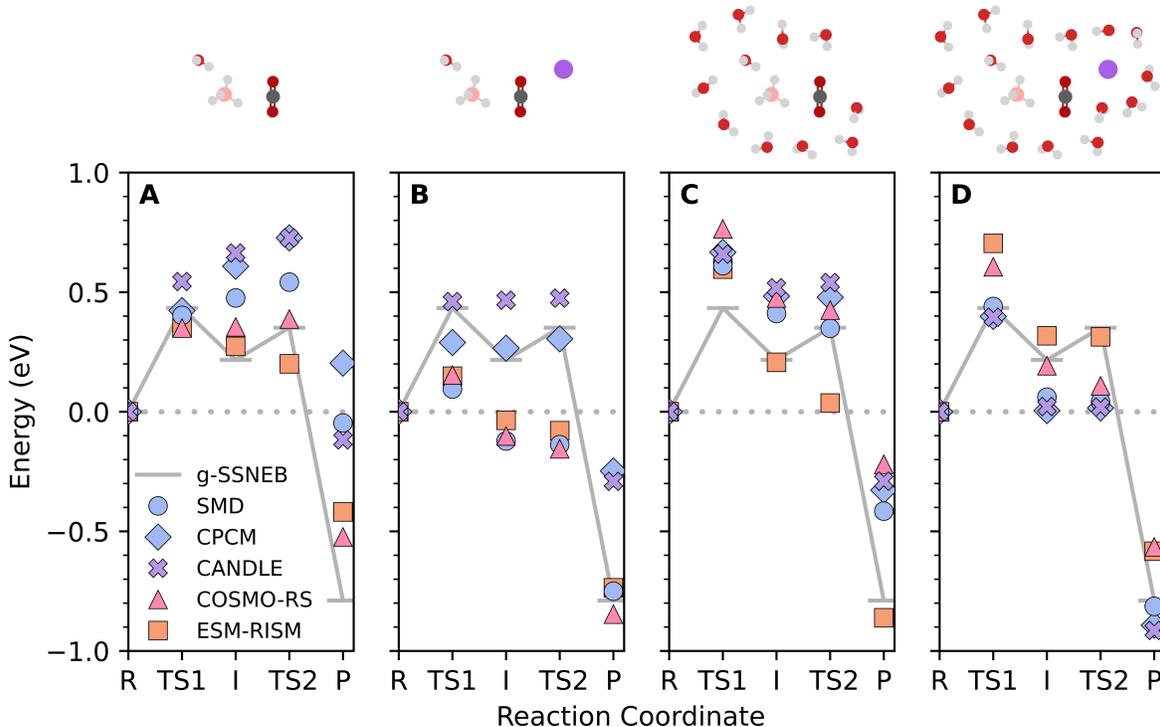


Figure 4: Comparison of solvated reaction energy profiles using SMD, CPCM, and CANDLE, COSMO-RS Fine, and ESM-RISM (points) compared to the fully explicit g-SSNEB reaction pathway using periodic PBE-PAW calculations (grey line) from Ref. 21. Data points represent the sum of the ω B97X-D3BJ/def2-TZVP gas-phase electronic energies and solvation energy contributions from the specified solvent model. Reaction energies are shown for (A) just the reacting molecules, (B) the reacting molecules with the Na^+ counterion, (C) the reacting molecules with the surrounding first solvent shell, and (D) the reacting molecules, the Na^+ counterion, and the first solvent shell including solvent molecules closest to the counterion.

Calculations for all g-SSNEB subsystems are shown in Figure 4. Illustrations of the g-SSNEB subsystem reactant state are shown above each graph, but the numbers of spectator water molecules are not equivalent to the actual system in 4C and 4D which are 21 and 37,

respectively. Figure 4A shows results using just reacting molecules (system charge of -1) solvated in a CSM. As was found previously,²¹ CSMs struggle to predict all local and bulk contributions; particularly as the reaction pathway proceeded forwards toward products as the hydride migrated from the middle of the cluster to the CO_2 for HCOO^- . SMD, CPCM, and CANDLE could not clearly identify the I state and resulted in qualitatively similar energy profiles. Furthermore, the error in the P state when using CPCM was so large that the overall reaction energy was predicted to be 0.2 eV uphill, while the explicit model (g-SSNEB) suggested that the overall process would be downhill by more than 0.8 eV. In contrast, COSMO-RS Fine and ESM-RISM predictions for the reaction pathway were more accurate (i.e., resembles the g-SSNEB explicit energies) than the other CSMs. While SMD, CPCM, and CANDLE found monotonically increasing reaction energies from TS1 to TS2, COSMO-RS Fine showed these three structures all resulted in about the same energy. We note that COSMO-RS (BP/TZVP/COSMO parametrization) followed the same qualitative reaction trend as COSMO-RS Fine, but the former was consistently higher in energy by almost 0.2 eV. ESM-RISM provided the best agreement with the reaction energies, and it was the only model to find the I state to be lower in energy than the TS1 state. However, ESM-RISM also resulted in the TS2 state being lower in energy. Based on this, one can see that CSMs should not be blindly trusted to predict reaction pathways with minimal information (i.e., no explicit solvents and counterion), but COSMO-RS and ESM-RISM are preferable.

Figure 4B shows results using the reacting molecules with the explicit Na^+ counterion present. Including the counterion causes all of the structures to be neutral, and this would be expected to significantly impact the results from the solvent models. In these cases, the relative energetics for the solvated reactions are lowered with respect to the R state. SMD, COSMO-RS Fine, and ESM-RISM all resulted in similar energies, and these models rather accurately reproduce the overall thermodynamics for this process. CPCM and CANDLE now do a better job recreating the relative energies of TS1, I, and TS2, but the energies of

the P states are only slightly lowered with respect to data in Figure 4A. One can see that the SMD, COSMO-RS Fine, and ESM-RISM do an excellent job recreating the local and bulk solvation energy contributions for the R and P states, but they all result in overly stabilized TS1, I, and TS2 states. In general, we see that adding just a counterion into a system with any solvent model does not systematically improve all of the states in the reaction pathway.

Figure 4C shows results when the reacting molecules are surrounded by an explicit solvent shell without a counterion. As with Figure 4A, all the system charges here are -1 . Interestingly, almost all the solvent models resulted in systematically destabilized reaction states except for ESM-RISM. ESM-RISM is the only case that accurately modeled the P state (as well as the I state), but the energies for TS1 and TS2 follow a similar qualitative trend as seen in Figures 4A and 4B. From this, we conclude that adding the first solvent shell (but not a counterion) can improve CSM predictions, but not always.

Figure 4D shows results when the reacting molecules are surrounded by an explicit solvent shell along with a counterion and its nearby solvent molecules. Thus, the solvent models here are only treating bulk solvent effects. The salient points here are that all solvent models now at least qualitatively capture the same effects as the fully explicit solvent model. Another interesting point is that the overall reaction energies are more accurately calculated using other CSMs rather than COSMO-RS Fine and ESM-RISM. Some CSMs can adequately perform relative energetics for molecular clusters, but the accuracies in the solvation energies can vary by about ± 0.2 eV when modeling transition states for these hydride migration processes.

We show these data as clear evidence that there will be at least a moderate degree of uncertainty (estimated to be about 0.2 to 0.3 eV) in static calculations representing any reaction state treated with any solvent model, *even when error cancellations across relative energetics across the same model are accounted for*. In general, all solvent models exhibited difficulties when characterizing the the same TS2 state as found in the explicit g-SSNEB model. We note that this energy barrier is quite small (0.13 eV in the g-SSNEB calculation

and only 0.05 eV in the BOMD PMF simulation) and would make it non-existent.

The system representation can be expected to significantly impact reaction energy profiles for solvated reactions. Minimal systems (i.e., no counterion or solvent shells) prove challenging for CSMs; some do perform moderately well, but are generally unreliable. Some (e.g., ESM-RISM) perform better than others; however, its performance could be due to explicitly accounting for specific molecular species in the reaction through the LJ parameters. From a pragmatic standpoint, it appears that the CSMs are most reliable when local solvent and counterion effects are treated explicitly. Thus, both are intriguing options for work in catalysis applications for renewable energy in the future, particularly since both are very promising for studying mixed and ionic solvent systems as well.⁹⁰

Continuum solvent model functional dependence

Many CSMs are parameterized to efficiently predict experimental solvation energies based on electronic densities from relatively low levels of theory and small basis sets.⁹¹ However, such methods are usually insufficient for modeling gas-phase reactions and could bring errors larger than 0.2 eV.⁹² This is why the current conventional wisdom is to use $E_{\text{liq}} - E_{\text{gas}}$ for solvation energy predictions with low levels of theory but highly accurate gas-phase energies. However, there is no guarantee that SCF errors in solvation energy calculations will always cancel. Figure 5 represents the span of possible solvated reaction energies when using accurate $\omega\text{B97X-D3BJ}/\text{def2-TZVP}$ gas-phase electronic energies with different CSM predictions from various levels of theory. Data for all g-SSNEB subsystems solvation energy predictions with respect to its functional are provided in the SI (Figure S8).

There appears to be only a small sensitivity on the underlying QM approach when using SMD and CPCM models with just the reacting atoms (5A). However, differences between PBE and ωB97X becomes much larger (0.5 eV) when the Na^+ counterion is included. In general, the presence of the counterion (Figures 5B and 5D) increases instabilities (relative to Figures 5A and 5C, respectively) of solvation energy predictions. However, since sensitivities

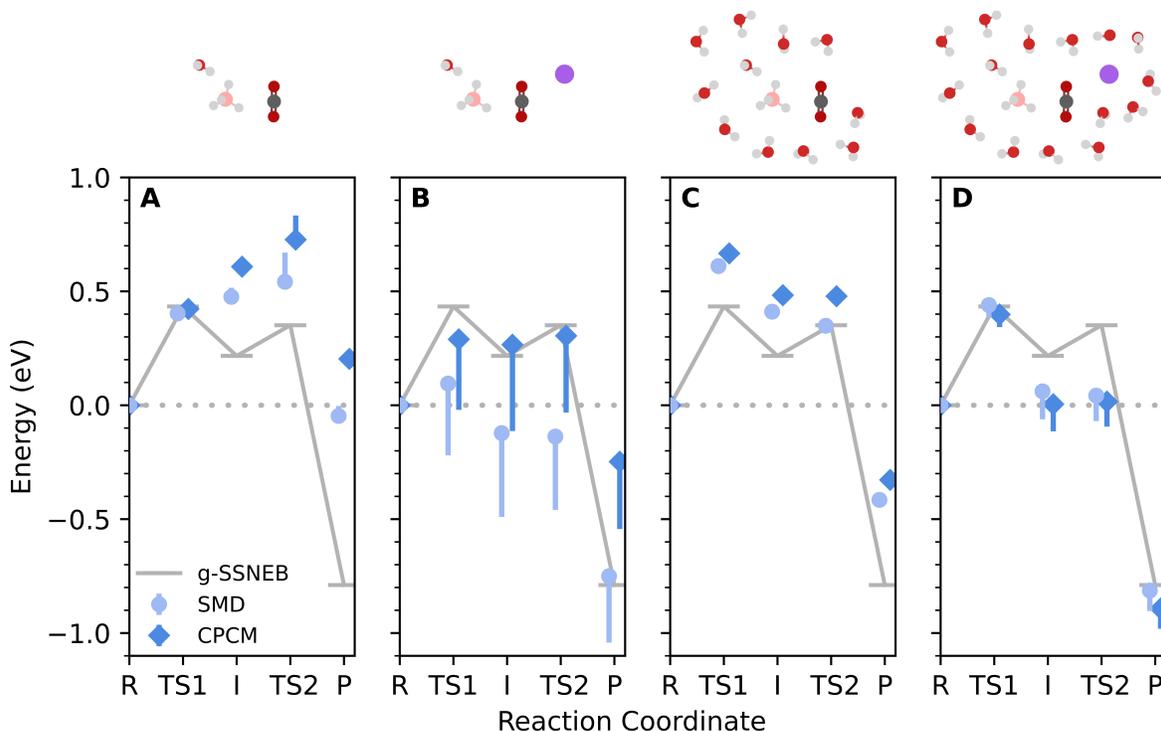


Figure 5: Sensitivities of SMD and CPCM to DFT functionals (PBE, BP86, B3LYP, and ω B97X). Each point represents ω B97X-D3BJ/def2-TZVP energies as shown in Figure 4. Vertical lines represent the span of energies possible when using solvation energies predicted with different levels of theory. Energies are shown for the (A) reacting molecules, (B) reacting molecules and counterion, (C) reacting molecules and solvent shell, (D) and reacting molecules, counterion, and solvent shell.

decrease from Figure 5B to 5D, it appears that the explicit solvent shell brings a stabilizing effect to the underlying electronic density treated by the solvent model. To probe this, molecular charges using the charges from electrostatic potentials using a grid-based method (CHELPG) scheme on the BH_3 , CO_2 , H_2O , Na^+ , and hydride (H) fragments using PBE and ω B97X are shown for gas (Figure 6A) and liquid phase (Figure 6B). Gas-phase molecular charges for all g-SSNEB subsystems are shown in the SI (Figures S9 and S10). There are significant CHELPG charge deviations for BH_3 and Na in the gas phase (Figure 6A) that are not present in the liquid phase (Figure 6B). This shows that the consistently large span of values shown by vertical lines shown in Figure 5B are due to errors when modeling the R state.

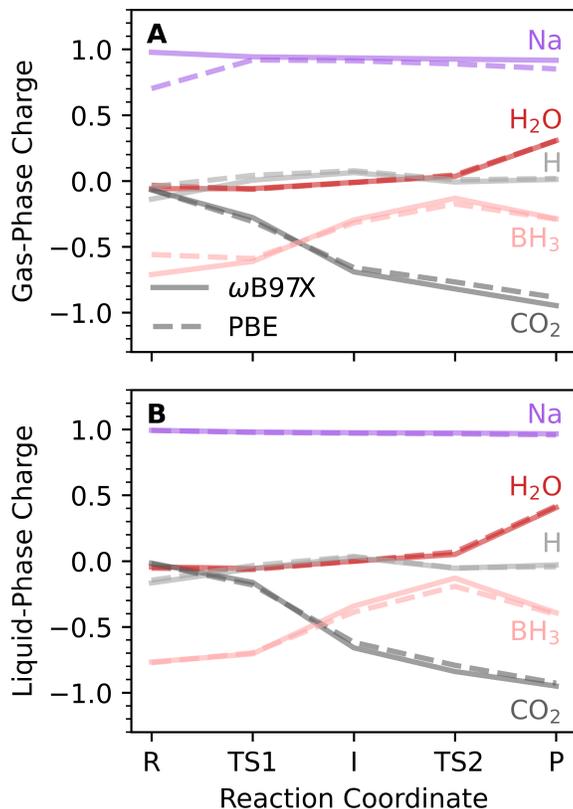


Figure 6: Molecular charges (from CHELPG) for reactants and counterion during the hydride transfer in (A) gas and (B) liquid phase (water as described CPCM). Charges from electron structure calculations using ω B97X is shown with solid lines while PBE is with dashes. Significant charge deviations in Na and BH₃ are observed in the gas phase reactant state, but they become corrected when using the CPCM solvent model.

A thermodynamic cycle can be used to decompose energy contributions into gas- and liquid-phase binding energies and solvation energies of the two chemical species. The gas-phase binding energy of Na⁺ to the reactants is significantly higher (by -0.371 eV) than ω B97X; indicating the solvation energy errors are primarily due to DFT electronic structure contributions in PBE calculations. An analysis of the thermodynamic cycle is presented in the SI (Figure S1 and Table S2).

Charge-separated states are notoriously challenging to model with DFT.⁹³ GGA functionals are known to exhibit substantial self-interaction errors while hybrid DFT functionals have less. By distancing Na⁺ from BH₄⁻, we probed the behavior of PBE and ω B97X functionals when treating the charge separation. As the distance between the two charges becomes

greater, the charges become primarily electrostatic in nature. We added 1 Å increments to the Na⁺ Cartesian coordinates and recalculated the molecular charges. Figure 7A shows a significant difference in the two charged species. At the initial separation of 7.5 Å, BH₄⁻ and

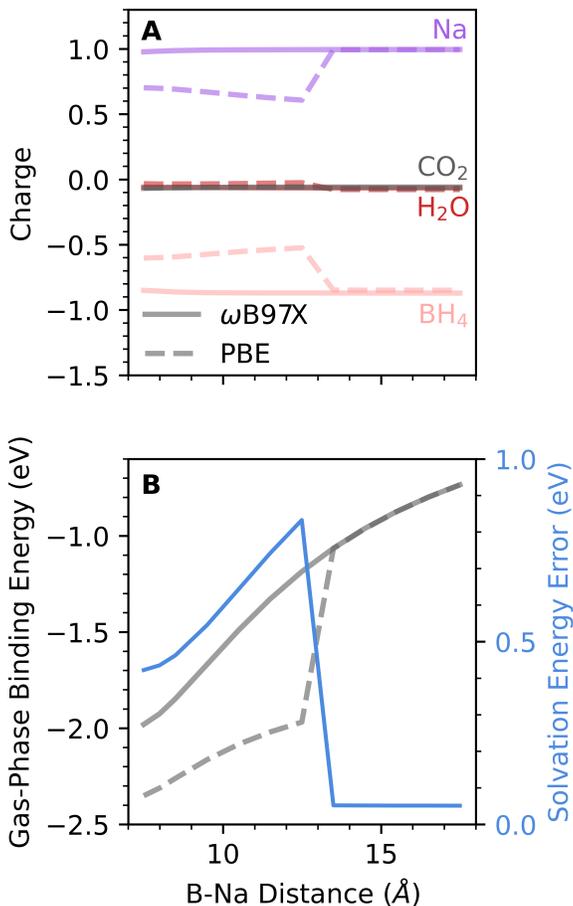


Figure 7: (A) CHELPG atomic charges for chemical species with respect to B-Na distance. (B) Large errors in PBE atomic charges cause significant differences in gas-phase binding energy due to self-interaction error. This carries over into the solvation free energy error as shown on the secondary y axis.

Na⁺ molecular charges are around 0.26 less in PBE than ωB97X. As the distance between BH₄⁻ and Na⁺ increases, the charge differences grow until they converge to the same value at a distance of 13.5 Å.

The consequences of these errors are shown in Figure 7B. The gas-phase binding energy error grows larger as the counterion is placed farther away, and this is carried over to solvation energy predictions. Once the counterion is far enough way (13.5 Å) the charges become

minimally interacting and PBE no longer overpredicts solvation energies by over 0.5 eV. This trend is actually analogous to results found by Carter and coworkers, where DFT functionals containing less than a threshold amount of exact exchange resulted in overly delocalized electron densities and qualitatively incorrect electronic states.⁹⁴ As a result, hybrid functionals are generally more reliable at avoiding self-interaction errors that can indirectly impact solvation energy predictions.

Single-ended reaction exploration

Computationally modeling the reaction using g-SSNEB with 70 explicit water molecules requires substantial computational resources. Single-ended GSM calculations with a few explicit solvent molecules, for example, can explore reaction space and provide preliminary results before embarking on fully explicit studies. Reaction predictions presented in this section represent the recommended technique for modeling a solvated reaction mechanism using a CSM: all structures are optimized and confirmed to be either a minimum or saddle point. The methods section above briefly mentions a multistep procedure (explained in more detail elsewhere¹³) to automate reaction discovery with minimal bias. All iterations of driving coordinates (e.g., breaking one of the B–H bonds and forming a H–C bond) are needed to completely sample pathways from an initial structure. Final structures are thus minimally biased and they require no additional computational effort to determine.

A variety of products were observed from single-ended GSM calculations of the same reactant state. Figure 8 shows two pathways that resulted in formate (the same product found by the g-SSNEB calculation). GSM pathways resulting in formic acid are in the SI. Note that formate was only observed in GSM calculations when the Na⁺ counterion was present. Predicted barrier heights and thermodynamics for the two reactions span a wide range of values in Figure 8A. CANDLE predicts the same rate-limiting energy barrier as the explicit g-SSNEB pathway did, while the COSMO-RS Fine and ESM-RISM predictions differ by 1 eV. However, adding a single spectator water molecule in the single-ended GSM

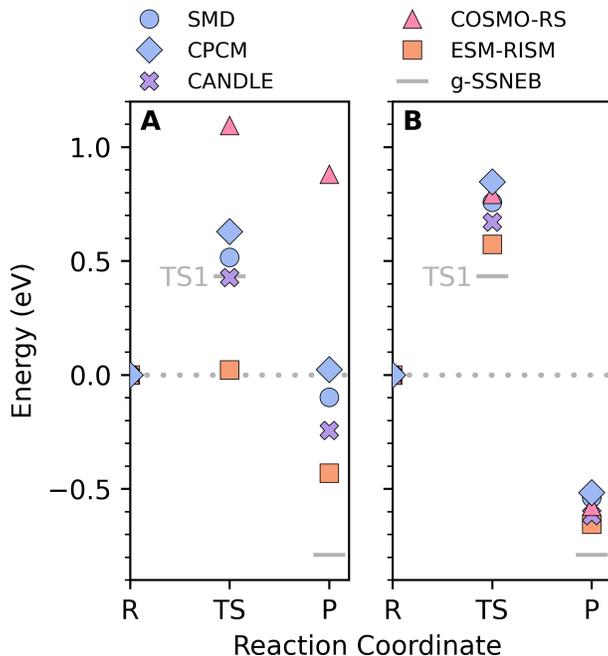


Figure 8: Single-ended GSM pathways that result in formate ($\text{BH}_3\text{OH}_2 + \text{HCOO}^-$) with (A) one and (B) two spectator water molecules. Both systems include a Na^+ counterion. Points are $\omega\text{B97X-D3BJ}/\text{def2-TZVP}$ gas-phase electronic energies with solvation energy contributions from the respective model. COSMO-RS is the FINE parameterization. The rate-limiting step for the g-SSNEB pathway (TS1) is 0.434 eV and the product at -0.789 eV.

(Figure 8B) appears to dramatically desensitize how different approaches model this system.

These single-ended GSM pathways present their own challenges. Specifically, single-ended GSM only produces pathways based on the atomic degrees of freedom available in the starting system. Thus, the two-step mechanism might be easily missed with a small cluster but more readily found with an explicit solvation procedure—for high throughput screening purposes this point may be inconsequential. Also, GSM can be useful for quick assessments of pathway viabilities as well as generate useful collective variables that are transferable to explicitly solvated dynamic simulations. For example, many of the pathways involve the system rearranging to better orient itself for a hydride transfer, and these pathways can be automatically found without the need for costly dynamics simulations. A potential downside of this approach is that the most robust use of GSM requires starting from a globally optimized structure. This can be difficult when force field parameters for more

unusual species (such as BH_4^-) may not readily be available, and standard Monte Carlo procedures would need to be used with QM (or semiempirical QM) calculations.

Conclusions

This work aims to show that modeling solvent effects on reaction mechanisms is a matter of calculated risks. Dynamic simulations involving many solvent molecules offers predictive confidence, but these are not always computationally feasible. The cost reduction of classic force fields in QM/MM simulations should continue to be considered, and factors such as counterion involvement in a fundamental charge migration in aqueous phase do not appear to be important. When dynamics simulations are not feasible, CSMs are much less computationally expensive, but they require many careful considerations.

In accordance with conventional wisdom, CSMs can be quite unreliable unless there are explicit solvent shells present to treat local solvent effects. A crucial aspect to look for in cases that may not have enough explicit solvent treatments is the degree that self-interaction errors are manifest, particularly in charge-separated states. Here, these are best treated using hybrid DFT (which may not be easily available in periodic boundary calculation) and/or with the use of suitably modeled explicit solvent shells surrounding the reacting molecules. Alternatively, COSMO-RS and ESM-RISM, are quite promising for predicting solvated reaction energies that implicitly account for interactions arising from local solvent and counterion contributions.

Data and conclusions drawn here are not necessarily valid for any reaction. However, one should be aware that uncertainties of 0.2 eV are possible regardless of the CSM used when studying a reaction mechanism, even in cases where one would expect significant error cancellation. For precise and accurate predictions of homogeneous and heterogeneous catalysis, we recommend more attempts to use explicit modeling whenever possible. These can benefit from efficient explorations of reaction steps using single-ended chain-of-states

methods to screen/preliminary evaluate pathways and seek useful collective variables for dynamics simulations in complex environments.

Supporting Information Available

The following files are available free of charge.

- A supporting information document with figures and tables of data mentioned in the main text.
- A repository of all QM/MM MD trajectories, output files, Python scripts for data analysis and figures, and CSV files (doi: 10.5281/zenodo.4336730).
- solvation-procedures-assessment.zip: contains XYZ and CSV files with a PDF of figures and tables.

Acknowledgement

We thank Mariana Rossi and Yair Litman for helpful discussions and diagnosing for us whether we should look into quantum nuclear effects. A.M., T.C., and J.K acknowledge support from the R. K. Mellon Foundation and the U.S. National Science Foundation (Grant Nos. CBET-1653392 and CBET-1705592). M.O. acknowledges partial financial support from MEXT as “Program for Promoting Researches on the Supercomputer Fugaku” (Fugaku Battery & Fuel Cell Project), grant number JPMXP1020200301. Some parts of computations were performed using the supercomputers of Information Technology Center at the University of Tokyo and Flow provided by Nagoya University.

References

- (1) BP plc, Statistical Review of World Energy. <https://bp.com/statisticalreview> (accessed Nov 24, 2020).
- (2) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. How enzymes work: analysis by modern rate theory and computer simulations. *Science* **2004**, *303*, 186–195.
- (3) Evans, D. H. One-electron and two-electron transfers in electrochemistry and homogeneous solution reactions. *Chem. Rev.* **2008**, *108*, 2113–2144.
- (4) Skubi, K. L.; Blum, T. R.; Yoon, T. P. Dual catalysis strategies in photochemical synthesis. *Chem. Rev.* **2016**, *116*, 10035–10074.
- (5) Markovitch, O.; Chen, H.; Izvekov, S.; Paesani, F.; Voth, G. A.; Agmon, N. Special pair dance and partner selection: Elementary steps in proton transport in liquid water. *J. Phys. Chem. B* **2008**, *112*, 9456–9466.
- (6) Kirchner, B.; di Dio, P. J.; Hutter, J. In *Multiscale Molecular Methods in Applied Chemistry*; Kirchner, B., Vrabec, J., Eds.; Springer: Berlin, Heidelberg, 2011; pp 109–153.
- (7) Acevedo, O.; Jorgensen, W. L. Advances in quantum and molecular mechanical (QM/MM) simulations for organic and enzymatic reactions. *Acc. Chem. Res.* **2010**, *43*, 142–151.
- (8) Klamt, A.; Eckert, F.; Arlt, W. COSMO-RS: an alternative to simulation for calculating thermodynamic properties of liquid mixtures. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1*, 101–122.
- (9) Barone, V.; Cossi, M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.

- (10) Nishihara, S.; Otani, M. Hybrid solvation models for bulk, interface, and membrane: Reference interaction site methods coupled with density functional theory. *Phys. Rev. B* **2017**, *96*, 115429.
- (11) Sundararaman, R.; Goddard III, W. A. The Charge-Asymmetric Nonlocally-Determined Local-Electric (CANDLE) Solvation Model. *J. Chem. Phys.* **2015**, *142*, 064107.
- (12) Sunoj, R. B.; Anand, M. Microsolvated transition state models for improved insight into chemical properties and reaction mechanisms. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12715–12736.
- (13) Basdogan, Y.; Keith, J. A. A paramedic treatment for modeling explicitly solvated chemical reaction mechanisms. *Chem. Sci.* **2018**, *9*, 5341–5346.
- (14) Pliego Jr, J. R.; Riveros, J. M. Hybrid discrete-continuum solvation methods. *WIREs Comput. Mol. Sci.* **2020**, *10*, e1440.
- (15) Plata, R. E.; Singleton, D. A. A case study of the mechanism of alcohol-mediated Morita Baylis–Hillman reactions. The importance of experimental observations. *J. Am. Chem. Soc.* **2015**, *137*, 3811–3826.
- (16) Schwarz, K.; Sundararaman, R. The electrochemical interface in first-principles calculations. *Surf. Sci. Rep.* **2020**, *75*, 100492.
- (17) Artz, J.; Müller, T. E.; Thenert, K.; Kleinekorte, J.; Meys, R.; Sternberg, A.; Bardow, A.; Leitner, W. Sustainable conversion of carbon dioxide: An integrated review of catalysis and life cycle assessment. *Chem. Rev.* **2018**, *118*, 434–504.
- (18) Song, Q.-W.; Zhou, Z.-H.; He, L.-N. Efficient, selective and sustainable catalysis of carbon dioxide. *Green Chem.* **2017**, *19*, 3707–3728.

- (19) Grice, K. A.; Groenenboom, M. C.; Manuel, J. D. A.; Sovereign, M. A.; Keith, J. A. Examining the selectivity of borohydride for carbon dioxide and bicarbonate reduction in protic conditions. *Fuel* **2015**, *150*, 139–145.
- (20) Li, P.; Henkelman, G.; Keith, J. A.; Johnson, J. K. Elucidation of aqueous solvent-mediated hydrogen-transfer reactions by ab initio molecular dynamics and nudged elastic-band studies of NaBH₄ hydrolysis. *J. Phys. Chem. C* **2014**, *118*, 21385–21399.
- (21) Groenenboom, M. C.; Keith, J. A. Explicitly unraveling the roles of counterions, solvent molecules, and electron correlation in solution phase reaction pathways. *J. Phys. Chem. B* **2016**, *120*, 10797–10807.
- (22) Groenenboom, M. C.; Keith, J. A. Quantum chemical analyses of BH₄⁻ and BH₃OH⁻ hydride transfers to CO₂ in aqueous solution with potentials of mean force. *ChemPhysChem* **2017**, *18*, 3148–3152.
- (23) Sheppard, D.; Xiao, P.; Chemelewski, W.; Johnson, D. D.; Henkelman, G. A generalized solid-state nudged elastic band method. *J. Chem. Phys.* **2012**, *136*, 074103.
- (24) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
- (25) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.
- (26) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **2020**, e01493.

- (27) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. et al. General atomic and molecular electronic structure system. *J. Comp. Chem.* **1993**, *14*, 1347–1363.
- (28) Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.
- (29) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (30) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112*, 8910–8922.
- (31) Grossfield, A. WHAM: the weighted histogram analysis method, version 2.0.4. http://membrane.urmc.rochester.edu/wordpress/?page_id=126.
- (32) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (33) Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **1994**, *50*, 17953.
- (34) Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. b* **1999**, *59*, 1758.
- (35) Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.
- (36) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (37) O’boyle, N. M.; Tenderholt, A. L.; Langner, K. M. Cclib: a library for package-independent computational chemistry algorithms. *J. Comput. Chem.* **2008**, *29*, 839–845.

- (38) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098.
- (39) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822.
- (40) Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
- (41) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785.
- (42) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (43) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys.* **2016**, *144*, 024109.
- (44) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (45) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (46) Zhong, S.; Barnes, E. C.; Petersson, G. A. Uniformly convergent n -tuple- ζ augmented polarized (n ZaP) basis sets for complete basis set extrapolations. I. Self-consistent field energies. *J. Chem. Phys.* **2008**, *129*, 184116.

- (47) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- (48) Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (49) Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning Jr, T. H.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg. *Theor. Chem. Acc.* **2011**, *128*, 69–82.
- (50) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.
- (51) Neese, F.; Valeev, E. F. Revisiting the atomic natural orbital approach for basis sets: Robust systematic basis sets for explicitly correlated and conventional correlated ab initio methods? *J. Chem. Theory Comput.* **2011**, *7*, 33–43.
- (52) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (53) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (54) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- (55) Grimme, S.; Bannwarth, C.; Shushkov, P. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions

- of large molecular systems parametrized for all spd-block elements ($Z= 1-86$). *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.
- (56) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory and Comput.* **2019**, *15*, 1652–1671.
- (57) Ehlert, S.; Bannwarth, C.; Caldeweyher, E.; Pracht, P.; Shushkov, P.; Grimme, S. xtb, version 6.2.3. 2020; <https://doi.org/10.5281/zenodo.3712016>.
- (58) Scalmani, G.; Frisch, M. J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* **2010**, *132*, 114110.
- (59) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian 16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.
- (60) Sundararaman, R.; Letchworth-Weaver, K.; Schwarz, K.; Gunceler, D.; Ozhabes, Y.; Arias, T. A. JDFTx: Software for Joint Density-Functional Theory. *SoftwareX* **2017**, *6*, 278–284.
- (61) Sundararaman, R.; Schwarz, K.; Letchworth-Weaver, K.; Arias, T. A. Spicing up continuum solvation models with SaLSA: the spherically-averaged liquid susceptibility ansatz. *J. Chem. Phys.* **2015**, *142*, 054102.
- (62) Dassault Systèmes, BIOVIA COSMOtherm, Release 2020. 2019.
- (63) Eckert, F.; Klamt, A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* **2002**, *48*, 369–385.
- (64) Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J. C.; Eckert, F. COSMO Implementation

- in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2187–2193.
- (65) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic structure calculations on workstation computers: The program system TURBOMOLE. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (66) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- (67) Reinisch, J.; Diedenhofen, M.; Wilcken, R.; Udvarhelyi, A.; Glöß, A. Benchmarking Different QM Levels for Usage with COSMO-RS. *J. Chem. Inf. Model.* **2019**, *59*, 4806–4813.
- (68) Andzelm, J.; Kölmel, C.; Klamt, A. Incorporation of solvent effects into density functional calculations of molecular energies and geometries. *J. Chem. Phys.* **1995**, *103*, 9312–9320.
- (69) Dassault Systèmes, BIOVIA COSMObase, Release 2020. 2019.
- (70) Laasonen, K.; Car, R.; Lee, C.; Vanderbilt, D. Implementation of ultrasoft pseudopotentials in ab initio molecular dynamics. *Phys. Rev. B* **1991**, *43*, 6796.
- (71) Laasonen, K.; Pasquarello, A.; Car, R.; Lee, C.; Vanderbilt, D. Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials. *Phys. Rev. B* **1993**, *47*, 10142.
- (72) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I. et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Condens. Matter Phys.* **2009**, *21*, 395502.
- (73) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Nardelli, M. B.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M. et al. Advanced capabilities for

- materials modelling with Quantum ESPRESSO. *J. Condens. Matter Phys.* **2017**, *29*, 465901.
- (74) Haruyama, J.; Ikeshoji, T.; Otani, M. Electrode potential from density functional theory calculations combined with implicit solvation theory. *Phys. Rev. Mater.* **2018**, *2*, 095801.
- (75) Kovalenko, A.; Hirata, F. Three-dimensional density profiles of water in contact with a solute of arbitrary shape: a RISM approach. *Chem. Phys. Lett.* **1998**, *290*, 237–244.
- (76) Smith, D. E.; Dang, L. X. Computer simulations of NaCl association in polarizable water. *J. Chem. Phys.* **1994**, *100*, 3757–3766.
- (77) Zhou, Y.; Yamaguchi, T.; Yoshida, K.; Fang, C.; Fang, Y.; Zhu, F. Structure of alkaline aqueous NaBH₄ solutions by X-ray scattering and empirical potential structure refinement. *J. Mol. Liq.* **2019**, *274*, 173–182.
- (78) Matsugami, M.; Yoshida, N.; Hirata, F. Theoretical characterization of the “ridge” in the supercritical region in the fluid phase diagram of water. *J. Chem. Phys.* **2014**, *140*, 104511.
- (79) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- (80) Zimmerman, P. M. Single-ended transition state finding with the growing string method. *J. Comput. Chem.* **2015**, *36*, 601–611.
- (81) Zimmerman, P. Reliable transition state searches integrated with the growing string method. *J. Chem. Theory Comput.* **2013**, *9*, 3043–3050.

- (82) Peters, B.; Heyden, A.; Bell, A. T.; Chakraborty, A. A growing string method for determining transition states: Comparison to the nudged elastic band and string methods. *J. Chem. Phys.* **2004**, *120*, 7877–7886.
- (83) Zimmerman, P. M. molecularGSM. 2016; <https://github.com/ZimmermanGroup/molecularGSM>.
- (84) Zhang, J.; Dolg, M. ABCcluster: the artificial bee colony algorithm for cluster global optimization. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24173–24181.
- (85) Zhang, J.; Dolg, M. Global optimization of clusters of rigid molecules using the artificial bee colony algorithm. *Phys. Chem. Chem. Phys.* **2016**, *18*, 3003–3010.
- (86) Milanese, J. M.; Provorse, M. R.; Alameda Jr, E.; Isborn, C. M. Convergence of computed aqueous absorption spectra with explicit quantum mechanical solvent. *J. Chem. Theory Comput.* **2017**, *13*, 2159–2171.
- (87) Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martinez, T. J. How large should the QM region be in QM/MM calculations? The case of catechol O-methyltransferase. *J Phys. Chem. B* **2016**, *120*, 11381–11394.
- (88) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.
- (89) Gillan, M. Quantum-classical crossover of the transition rate in the damped double well. *J. Phys. C: Solid State Phys.* **1987**, *20*, 3621.
- (90) Maldonado, A. M.; Basdogan, Y.; Berryman, J. T.; Rempe, S. B.; Keith, J. A. First-principles modeling of chemistry in mixed solvents: Where to go from here? *J. Chem. Phys.* **2020**, *152*, 130902.

- (91) Ho, J.; Klamt, A.; Coote, M. L. Comment on the correct use of continuum solvent models. *J. Phys. Chem. A* **2010**, *114*, 13442–13444.
- (92) Hodgson, J. L.; Namazian, M.; Bottle, S. E.; Coote, M. L. One-electron oxidation and reduction potentials of nitroxide antioxidants: A theoretical study. *J. Phys. Chem. A* **2007**, *111*, 13595–13605.
- (93) Solovyeva, A.; Pavanello, M.; Neugebauer, J. Describing long-range charge-separation processes with subsystem density-functional theory. *J. Chem. Phys.* **2014**, *140*, 164103.
- (94) Oyeyemi, V. B.; Keith, J. A.; Pavone, M.; Carter, E. A. Insufficient Hartree–Fock exchange in hybrid DFT functionals produces bent alkynyl radical structures. *J. Phys. Chem. Lett.* **2012**, *3*, 289–293.

TOC Graphic

