



FEG optimizations using dual-level approach II: Vibrational Spectroscopy in Solution

Carlos BISTAFA

Graduate School of Informatics, Nagoya University, Japan

5th CREST Workshop FY2018
Nagoya, March 5th, 2019

Previous presentations

- *"Microscopic Origin of Different Hydration Pattern of para-Nitrophenol and Its Anion: A Study Combining Multiconfigurational Calculations and the Free-Energy Gradient Method"*
Bistafa, Kitamura, Nagaoka, Canuto **JPCB** 122 (2018) 9202
- *"Theoretical electronic spectra dependent on pH conditions. A joint application of CpH-MD simulations and QM/MM"*
Bistafa, Kitamura, Nagaoka (manuscript in preparation)
- *"A Cost-Effective Method for Free-Energy Minimization in Complex Systems with Elaborated Ab Initio Potentials"*
Bistafa, Kitamura, Martins-Costa, Nagaoka, Ruiz-López, **JCTC** 14 (2018) 3262
- *"Vibrational Spectroscopy in Solution Through Perturbative ab initio Molecular Dynamics Simulations"*
Bistafa Kitamura, Martins-Costa, Nagaoka, Ruiz-López (to be submitted to **JCTC** soon)

FEG-FEP

Previous CREST WS

Using FEP equations, we can prove that the derivatives of the free energy at HL are related to the LL by:

$$\frac{\partial G_{HL}(q)}{\partial q} = \left\langle \frac{\partial V_{HL}(q;s)}{\partial q} w(q;s) \right\rangle_{LL}$$

Where w is a weighting factor expressed by:

$$w(q;s) = \frac{e^{-\beta\Delta V(q;s)}}{\left\langle e^{-\beta\Delta V(q;s)} \right\rangle_{LL}}$$

We use 10% of the configurations sampled at LL to calculate the values at HL.

JCTC
Journal of Chemical Theory and Computation

Article
Cost-Effective Method for Free-Energy Minimization in Complex Systems with Elaborated Ab Initio Potentials

Carlos Bistafa,¹ Yukichi Kitamura,² Marilía T. C. Martins-Costa,³ Masataka Nagaoka,^{4,5,6} and Manuel F. Ruiz-López^{4,7,8}

¹Department of Complex Systems Science, Graduate School of Informatics, Nagoya University, Chikusa Ku, Furo Cho, Nagoya, Aichi 4648601, Japan

²Laboratoire de Physique et Chimie Théoriques, UMR CNRS 7019, Faculté des Sciences et Technologies, Université de Lorraine, CNRS, BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France

³ESKCB, Kyoto University, Kyodai Katsura, Nishihyogo-ku, Kyoto 6158520, Japan

⁴Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Honmachi, Kawaguchi 3320012, Japan

⁵Future Value Creation Research Center, Graduate School of Informatics, Nagoya University, Chikusa Ku, Furo Cho, Nagoya, Aichi 4648601, Japan

⁶Supporting Information

ABSTRACT: We describe a method to locate stationary points in the free-energy hypersurface of complex molecular systems using high-level correlated *ab initio* potentials. In this work, we assume a combined QM/MM description of the system although generalization to full *ab initio* potentials or other theoretical schemes is straightforward. The free-energy gradient (FEG) is obtained as the mean force acting on relevant nuclei using a dual level strategy. First, a statistical simulation is carried out using an appropriate, low-level quantum mechanical force-field. Free-energy perturbation (FEP) theory is then used to obtain the free-energy derivatives for the target, high-level quantum mechanical force-field. We show that this composite FEG-FEP approach is able to reproduce the results of a standard free-energy minimization procedure with high accuracy, while simultaneously allowing for a drastic reduction of both computational and wall-clock time. The method has been applied to study the structure of the water molecule in liquid water at the QCISD/aug-cc-pVTZ level of theory, using the sampling from QM/MM molecular dynamics simulations at the B3LYP/6-311+G(d,p) level. The obtained values for the geometrical parameters and for the dipole moment of the water molecule are within the experimental error, and they also display an excellent agreement when compared to other theoretical estimations. The developed methodology represents therefore an important step toward the accurate determination of the mechanism, kinetics, and thermodynamic properties of processes in solution, in enzymes, and in other disordered chemical systems using state-of-the-art *ab initio* potentials.

1. INTRODUCTION

The theoretical study of chemical reactions in solution and in complex molecular environments often requires the knowledge of the solute's geometry at the stationary points of the free-energy landscape of the system (reactants, products, intermediates, transition structures). Since the pioneer applications of implicit continuum models to describe reaction paths in solution,¹⁻³ thousands of new studies have been carried out using those models, providing valuable insights into solvation effects on the kinetics, thermodynamics, and/or selectivity of chemical processes. In continuum models (for further details see reviews reported before^{4,5}), it is straightforward to compute the derivatives of the free energy of solvation with respect to the solute's parameters^{6,7} and therefore to locate the relevant stationary points. However, due to crude approximations, implicit models are not always

sufficiently accurate, and large efforts have been made in the last years to develop more elaborated models.

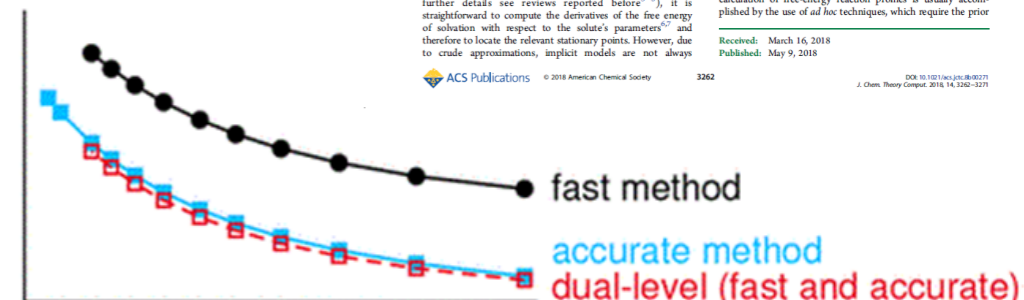
In this line, Molecular Dynamics (MD) using QM/MM (Quantum Mechanics/Molecular Mechanics) force-fields has become one of the most popular and successful techniques to study complex molecular systems.⁸ The price to be paid is a much higher computational cost, and hence most applications assume semiempirical or density functional theory QM methods and MD sampling limited to a few tens or hundreds of picoseconds (though recent algorithms⁹ have been allowed to reach multisecond time scales). In MD simulations, the calculation of free-energy reaction profiles is usually accomplished by the use of *ad hoc* techniques, which require the prior

Received: March 16, 2018
Published: May 9, 2018

ACS Publications | © 2018 American Chemical Society

3262

DOI: 10.1021/acs.jctc.8b00271
J. Chem. Theory Comput. 2018, 14, 3262–3271



Results: Optimized Structure of Water Molecule in Aqueous Solution

Previous CREST WS

Method	Property	Gas	Liquid	Δ
B3LYP/6-311+G(d,p) FEG^a	d_{OH} (Å)	0.962	0.982	+0.020
	α (°)	105.1	105.2	+0.1
	μ (D)	2.16	2.91	+0.75
B3LYP::HF (FEG-FEP)	d_{OH} (Å)	0.962	0.982	+0.020
	α (°)	105.1	105.2	+0.1
	μ (D)	2.16	2.95	+0.79
QCISD::B3LYP (FEG-FEP)	d_{OH} (Å)	0.959	0.979	+0.020
	α (°)	104.4	105.5	+1.1
	μ (D)	1.86	2.69	+0.83
Experimental values	d_{OH} (Å)	0.957 ^b	0.970±0.005 ^c	+0.013
	α (°)	104.5 ^b	~106±1.7 ^c	+1.5
	μ (D)	1.85 ^d	2.6 ^e	+0.75

^a Standard FEG optimization;

^b Benedict et. al, *J. Chem. Phys.*, **24**, 1139 (1956);

^c Ichikawa et. al., *Mol. Phys.*, **73**, 79 (1991);

^d Clough et. al., *J. Chem. Phys.*, **59**, 2254 (1973);

^e Coulson et. al., *Proc. R. Soc. London., Ser. A*, **291**, 445 (1966).

Motivation

Can we also do that for the Hessian?

If we can, it would allow us to obtain results for frequencies (consequently, IR and Raman spectra), and also to study transition states.

Dual level calculations for Hessian (FEP-FEH)

By differentiating the FEG-FEP equation, we can prove that the Hessian of the free energy obtained at a High Level (HL) sampling is related to a low level (LL) sampling by:

$$\frac{\partial^2 G_{HL}(q)}{\partial q \partial q} = \left\langle \frac{\partial^2 V_{HL}(q;s)}{\partial q \partial q} \right\rangle_{HL} = \left\langle \frac{\partial^2 V_{HL}(q;s)}{\partial q \partial q} w(q;s) \right\rangle_{LL}$$

Where w is the same weighting factor we have obtained for the gradient case:

$$w(q;s) = \frac{e^{-\beta \Delta V(q;s)}}{\left\langle e^{-\beta \Delta V(q;s)} \right\rangle_{LL}}$$

We use 10% of the configurations sampled at LL to calculate the values at HL.

Computational Details

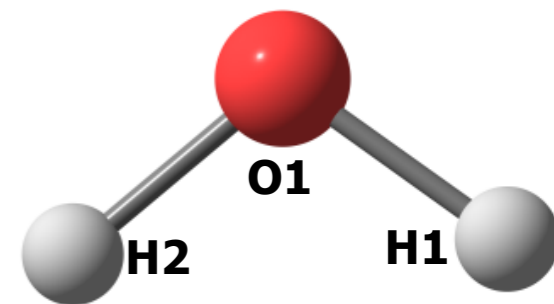
- **System: 1 QM water molecule + 592 TIP3P water molecules.**
(small, interesting and data is available for comparison)

- **Simulation details:**

B3LYP::HF -> QM/MM simulations performed at the HF/6-31G(d) level, with Hessian calculations corrected to B3LYP/6-311++G(d,p).

In order to validate our approach, usual QM/MM simulation and Hessian calculations at the B3LYP/6-311++G(d,p) level were also performed.

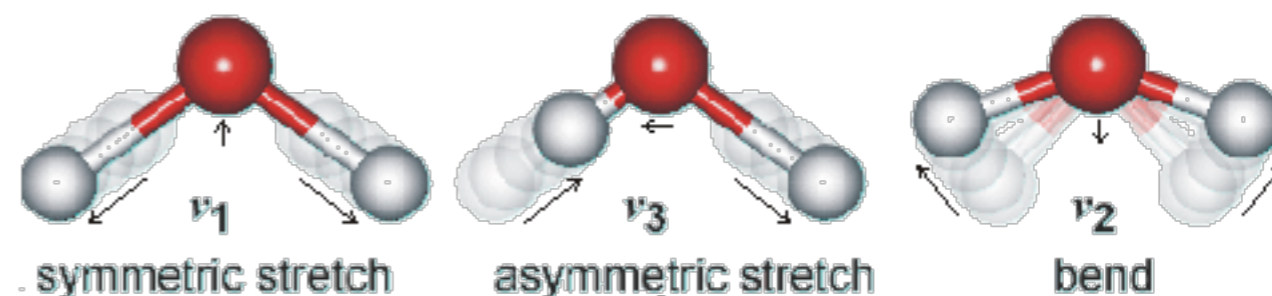
NVT, 300K, 40 ps, time-step of 1fs.



Results 1: Vibrational Frequencies in Gas Phase

Table 1: Theoretical and experimental vibrational frequencies of the water molecule in the gas phase (cm^{-1}). Calculated anharmonic frequencies are estimated using the second-order perturbative approach or from scaled harmonic frequencies (in parenthesis).

		Bend	Symm Stretch	Asymm Stretch
HF/6-31G(d)^a	Harmonic	1826	4069	4188
B3LYP/6-311+G(d,p)^a	Harmonic	1601	3817	3923
	Anharmonic	1555 (1539)	3646 (3668)	3739 (3770)
MP2/aug-cc-pVDZ^b	Harmonic	1631	3846	3972
	Anharmonic	1578	3674	3788
CAS/ANO[6s5p4d3f/5s4p3d]^c	Harmonic	1657	3849	3965
Experiment^d		1595	3657	3756



^a This work;

^b Georg & Canuto, *J. Phys. Chem. B*, **116**, 11247 (2012);

^c Nymand & Åstrand, *J. Phys. Chem. A*, **101**, 10039 (1997);

^d Benedict et. al., *J. Chem. Phys.*, **24**, 1139 (1956);

Results 2: Vibrational Frequencies in Aqueous Solution

Table 2: Calculated vibrational frequencies in the liquid phase and frequency shifts with respect to the gas phase of the water molecule using different methods (cm⁻¹). Anharmonic frequencies are obtained by scaling the calculated harmonic values by 0.961.

		Bend	Symm Stretch	Asymm Strech
FEP-FEH	Harmonic	1755	3501	3571
	Anharmonic	1687	3364	3432
	$\Delta(\text{Gas-Liquid})$	+148	-304	-338
FEH	Harmonic	1735	3514	3587
	Anharmonic	1667	3377	3447
	$\Delta(\text{Gas-Liquid})$	+128	-291	-323
Experiment^a		1650	2800-3700	
	$\Delta(\text{Gas-Liquid})$	~ +55	~ -250	

^a Bertie & Eysel, *Applied Spectroscopy*, **39**, 392 (1985);
Schmidt & Miki, *J. Phys. Chem. A*, **111**, 10119 (2007);
Eisenberg & Kauzmann, *The structure and properties of water*, Oxford University Press, NY, 1969;
Bertie & Lan, *Applied Spectroscopy*, **50**, 1047 (1996).

Theoretical Spectra

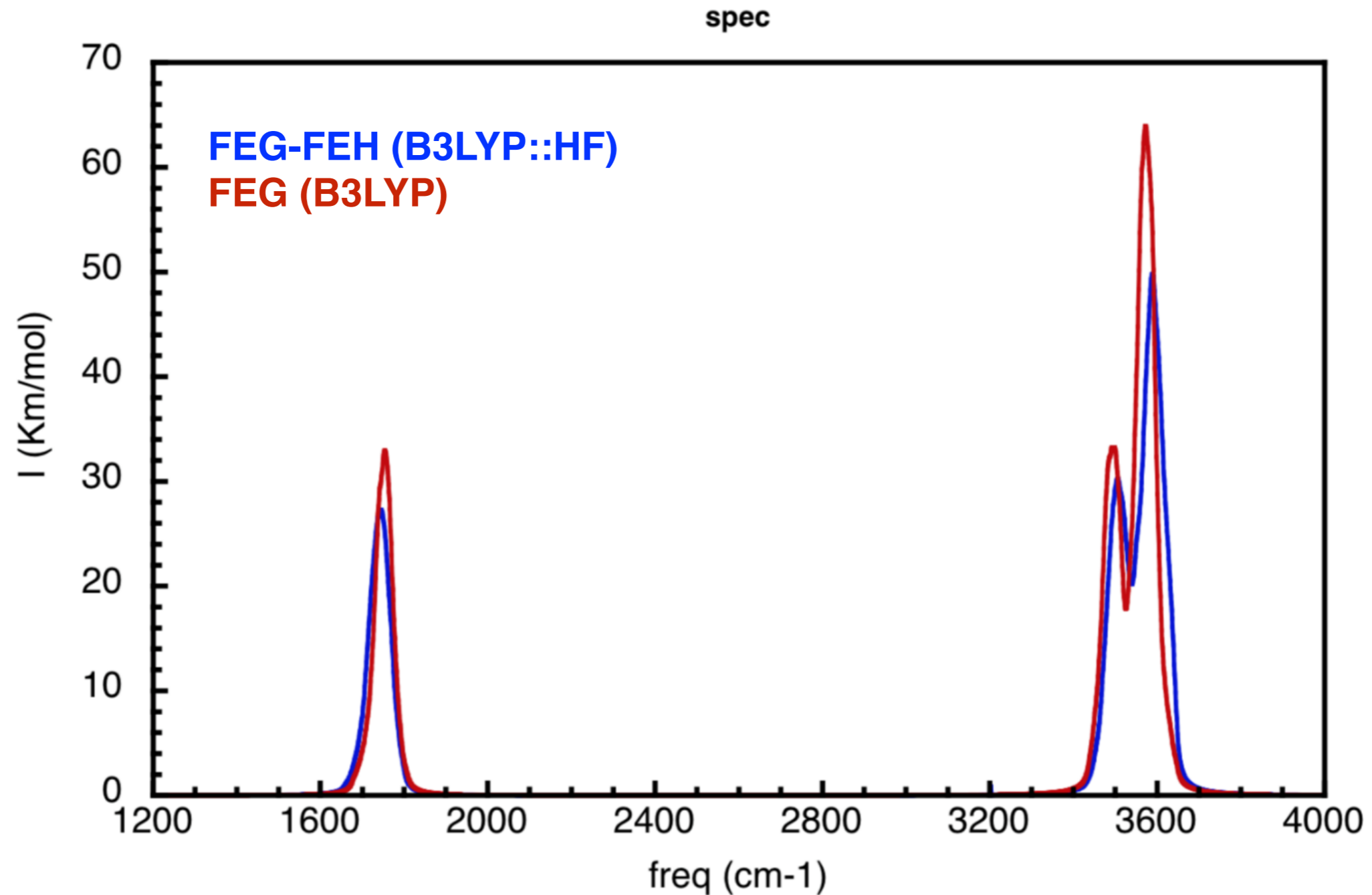


Figure 1: Theoretical spectra obtained by the convolution of a Lorentzian function over the frequencies and their respective intensity values that were calculated in both approaches.

Theoretical and Experimental Spectra

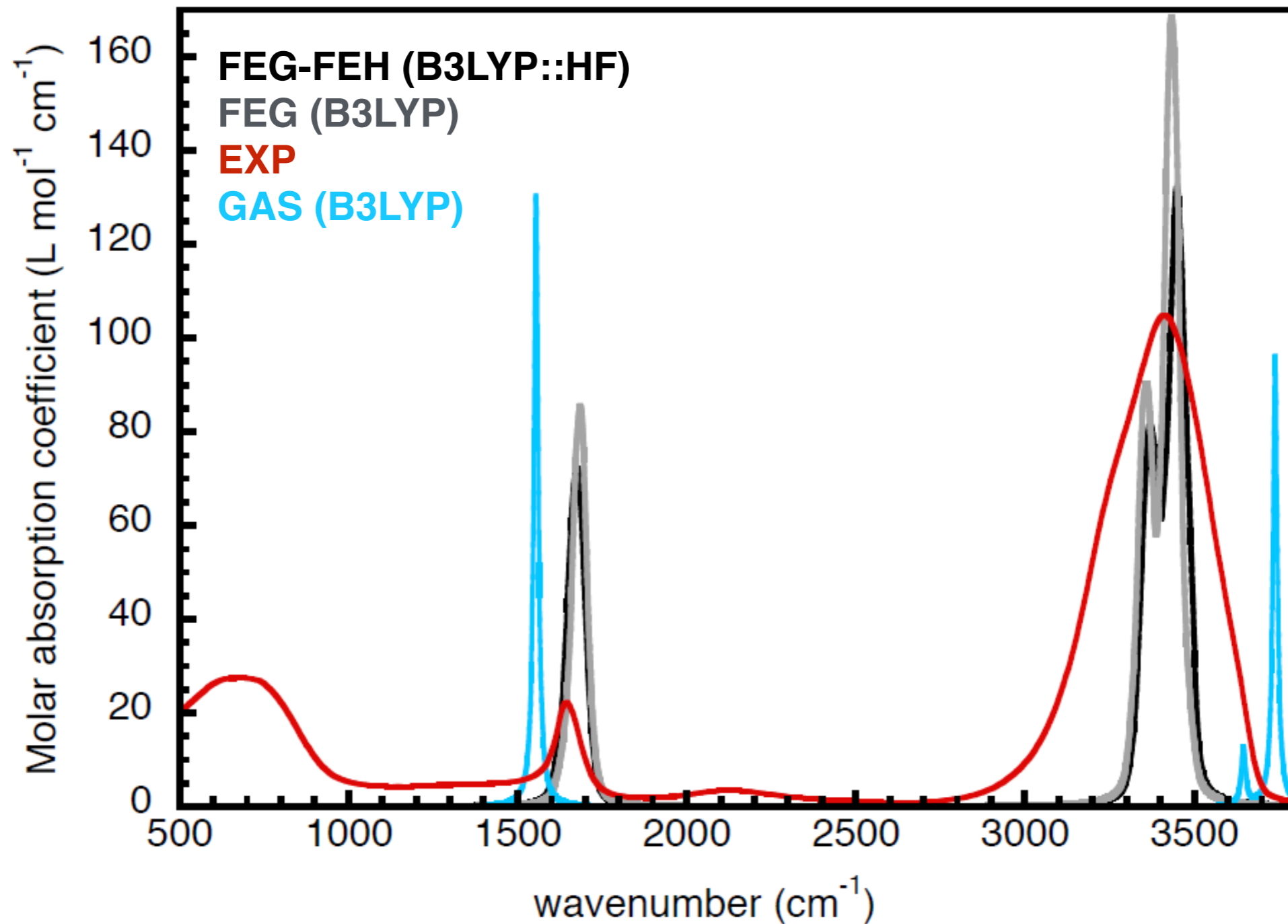


Figure 2: Theoretical and experimental infrared spectra of water molecule.

Conclusions

- We have shown that FEP methods combined with samplings from affordable QM/MM simulations are capable of reproducing results at high-levels of theory.
- As a test case, we have investigated the IR spectrum of the water molecule in liquid water showing that it can be properly reproduced by combining QM/MM simulations carried out at the simple HF/6-31G(d) level and perturbational corrections calculated at the B3LYP/6-311+G(d,p) level.

Conclusions

- From the computational point of view, the combination of QM/MM simulations and FEP corrections is quite efficient;
- The possibility to calculate accurate Hessian matrices, combined with the possibility of calculating accurate gradients developed in our previous work open a path for the study of chemical reactions at unprecedented levels of theory.