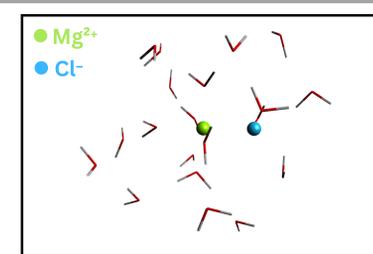


Investigating the Viability of Gaussian Process Regression to Produce a 1-Dimensional Free Energy Landscape of the Mg^{2+} - Cl^- Ions Association Process in an Aqueous Solution

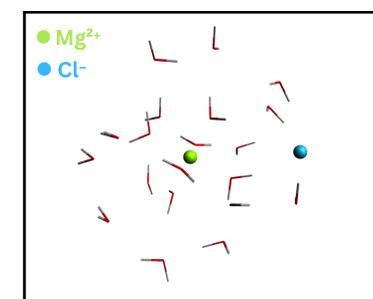
Sikharin Senjai, Advisor : Dr. Wasut Pornpatcharapong
Department of Chemistry, Faculty of Science, Chiang Mai University

ABSTRACT

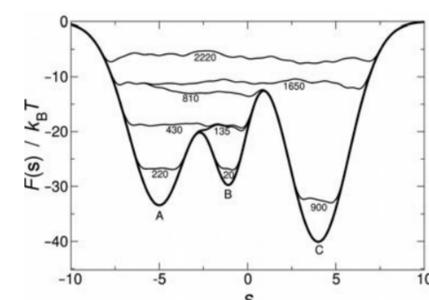
The free energy surface (FES) is crucial for identifying the relative stability of a system and can also elucidate the mechanism of chemical reactions through the minimum free energy path (MFEP). In this work, we explore the integration of Well-Tempered Metadynamics (WT-MTD) with Gaussian Process Regression (GPR) to efficiently reconstruct the free energy surface. This approach is applied to study the ion-pairing process of Mg^{2+} and Cl^- in an aqueous environment. The accuracy of the reconstructed free energy surface from Gaussian Process Regression (GPR) the results are compared with free energy profiles obtained from Umbrella Sampling (WHAM) and literature data. The reference profile is selected based on the most closely aligns with the literature results. Although applying Gaussian Process Regression (GPR) to a one-dimensional free energy surface might seem to add complexity to the Well-Tempered Metadynamics (WT-MTD) simulation, which already provide reasonable free energy estimates, this study demonstrates that Gaussian Process Regression (GPR) can accurately reconstruct the free energy surface, achieving high accuracy with relatively small training datasets.



contact ion-pairs (CIPs)



SSHIPs (Solvent-Shared Ion-Pairs)



INTRODUCTION

Free energy landscapes are crucial for understanding chemical processes, particularly those involving metastable states, requires an accurate representation of free energy landscapes. These landscapes are typically projected onto collective variable (CV) spaces to facilitate interpretation. However, traditional methods like umbrella sampling (US) are computationally expensive, especially in high-dimensional CV spaces. Metadynamics (MTD) and its well-tempered variant (WT-MTD) provide a more efficient approach by applying history-dependent bias potentials. While effective, these methods can yield rough energy landscapes that require post-processing. However, these methods can produce rough landscapes, requiring post-processing. Recent studies have suggested combining MTD/WT-MTD with machine learning techniques, such as Gaussian Process Regression (GPR), to reconstruct smooth free energy landscapes. This study explores the feasibility of GPR in reconstructing the free energy landscape of the Mg^{2+} - Cl^- ion pairing process in water, comparing it with US and WT-MTD. The objective is to evaluate GPR's accuracy, computational efficiency, and applicability to multidimensional systems, paving the way for enhanced free energy estimation in complex chemical processes.

OBJECTIVES

- Compare the free energy profiles obtained from umbrella sampling (US) and WT-MTD to identify the most reliable reference free energy landscape.
- Evaluate the performance of GPR in reconstructing smooth and accurate free energy landscapes using free energy derivatives derived from WT-MTD trajectories.

CONCLUSION

This study demonstrates that combining Well-Tempered Metadynamics (WT-MTD) and Gaussian Process Regression (GPR) enables efficient and accurate free energy landscape reconstruction. GPR effectively smooths WT-MTD-derived landscapes with minimal training data. This technique has high potential for reconstructing multidimensional free energy landscapes, where traditional methods like US and WT-MTD face scalability limitations. WHAM results showed discrepancies due to insufficient sampling, particularly in high-energy (barrier) regions.

METODOLOGY

System Preparation

- Construct a simulation box (Mg^{2+} , Cl^- , TIP3P water molecules).
- Energy minimization and 4 ns equilibration (NVT ensemble, 300K).



Umbrella Sampling (US) Simulation

- Define d_{Mg-Cl} as the collective variable (CV).
- Conduct simulations across 49 windows.
- Analyze with WHAM to generate reference free energy profile.



WT-MTD Simulation

- Perform 1D WT-MTD using OpenMM-PLUMED.
- Generate 10 ns and 50 ns trajectories to obtain biased free energy landscapes.
- Extract free energy derivatives from WT-MTD data.



GPR for Free Energy Reconstruction (<https://github.com/kulhanek/pmflib>)

- Compute unbiased free energy derivatives.
- Train GPR model with 5,000 and 25,000 data points.
- Optimize GPR hyperparameters (δ , σ , θ).



Validation and Computational Efficiency Optimization

- Compare GPR results with WT-MTD and WHAM landscapes.
- Evaluate accuracy using error norm analysis.

Metadynamics settings

Add a Gaussian every 1,000 steps
Gaussian height: 0.4 kcal/mol
Gaussian width: 0.1 Angstroms
Using Well-tempered MetaD with the value of $\Delta T = 8,500$ K
Using 1-CV: Distance between Mg and an oxygen atom of 1 water molecule

```
# RUNNING GPR
# TEST FOR ONE SET OF HYPERPARAMETERS

import time

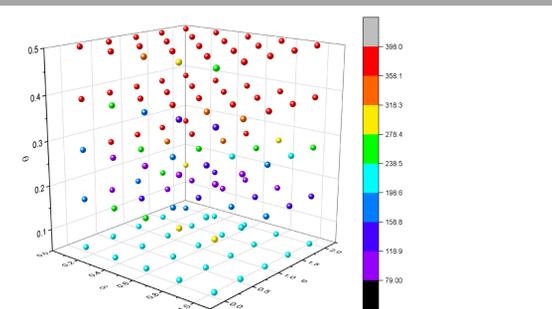
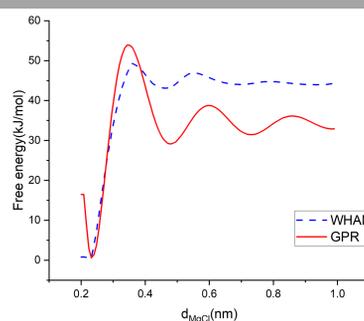
os.environ['OMP_NUM_THREADS'] = '8'

delta = 0.5
theta = 0.01
sigmader = 2.5
predgrid = 50
ndim = 1

t1 = time.time()

gpr_exec = './../program_gp_general/bin/gp_general'
index = -1
```

RESULTS AND DISCUSSION



Graph (Left). Comparison of the GPR-reconstructed free energy landscape from 5,000 - point training data of free energy gradients to the WHAM-constructed free energy surface.

Graph (Right). Effects of GPR hyperparameters on the error norm of the 1D free energy landscape based on the d_{Mg-Cl} CV relative to the WT-MTD reference surface for the 5,000-point training data.

REFERENCES

1. Mones, L.; Bernstein, N.; and Csányi, G. Exploration, sampling, and reconstruction of free energy surfaces with Gaussian process regression. *J. Chem. Theory Comput.* **2016**, *12*, 5100–5110
2. Barducci A.; Bonomi M.; and Parrinello M. Metadynamics© 2011 John Wiley & Sons, Ltd. *WIREs Comput Mol Sci*, **2011**, *1*, 826–843