Computational Chemistry Enabled by Quantum Computers

2026 NSF CSSI/CT/SCIPE PI Meeting
Washington, DC

September 13-15, 2026

Kenneth M. Merz, Jr.
Center of Computational Life Science
Cleveland Clinic and
Adjunct Professor of Chemistry
Department of Chemistry
Michigan State University

Baia del Silenzio, Sestri Levante, Italy

Acknowledgments

Cleveland Clinic Lerner Research Institute

Kennie Merz

Akhil Shajan
Danil Kaliakin
Fangchun Liang
Zhen Li
Milana Bazayeva
Susanta Das
Subhamoy Bhowmik

IBM

Mario Motta
Ella Fejer

Thaddeus Pellegrini
Javier Robledo Moreno
Abdullah Ash Saki
Kevin Sung
Ella Fejer
Robert Walkup
Seetharami Seelam

RIKEN

Mitsuhisa Sato
Seiji Yunoki

Yuichi Otsuka
Tomonori Shirakawa
Lukas Broers
Han Xu
Miwako Tsuji

NIH NSF

Ryo Wakizaka
Yukio Kawashima
Jun Doi
Toshinari Itoko
Hiroshi Horii

Hardware Access is Essential

On-premises IBM Quantum System One Eagle 127 qbits (unveiled March 2023)

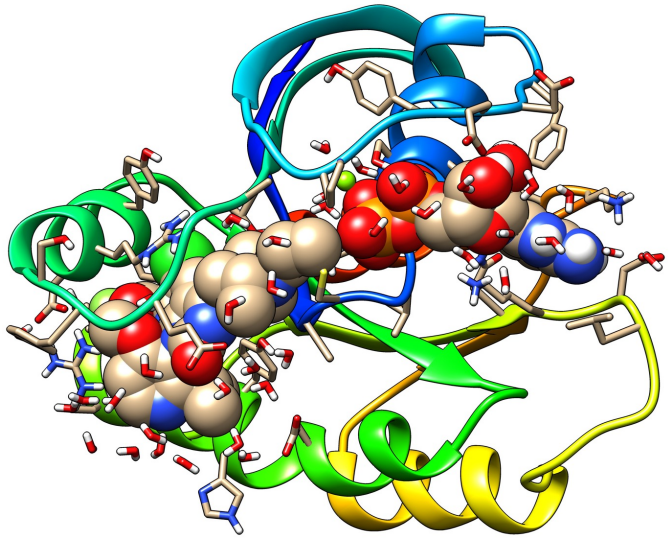
Upgraded to Heron r2 156 qbits (2026)

Integrated into a quantum centric supercomputing (QCSC) environment

Quantum System One is **dedicated to healthcare and life sciences** with the mission to create a HCLS innovation ecosystem with outreach and impact worldwide

What is the Nature of Your Problem?

Overarching goal: Routinely study large biologically relevant systems using Quantum Chemistry on Quantum Computers



- Intermolecular Interactions
  - Dispersion interactions
  - Hydrogen bonding

  See: https://www.nature.com/articles/s42005-025-02305-9
- Large systems
  - Drug-like molecules
  - Biomacromolecules

  See: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.5c00114
  See https://arxiv.org/abs/2605.01138
- Solvation
  - Continuum

  See: https://pubs.acs.org/doi/full/10.1021/acs.jpcb.5c01030
  - Explicit

  FEP in explicit solvent: https://arxiv.org/html/2506.20825v1
  FEP in protein-ligand system: https://arxiv.org/abs/2604.09857

Calderasib and GDP and Mg2+ in the KRAS G12C

Vacca, J. P.; Dorsey, B. D.; et al. *Proc Natl Acad Sci U S A* **1994**, *91*, 4096

Scaling in Quantum Chemistry
Fragment Linear-Scaling in Quantum Chemistry

VOLUME 66, NUMBER 11 PHYSICAL REVIEW LETTERS 18 MARCH 1991

Direct Calculation of Electron Density in Density-Functional Theory

Weitao Yang
Department of Chemistry, Duke University, Durham, North Carolina 27706
(Received 5 September 1990)

A new approach for the study of ground states of many-electron systems is developed via direct calculation of the density in density-functional theory. Not using the Kohn-Sham equations, the method divides a system into subsystems in physical space and determines the density for each subsystem. The method is demonstrated with calculations for the nitrogen molecule, which is divided into two atomic subsystems. We expect this approach to enable calculations for large molecules beyond the reach of conventional methods.

PACS numbers: 31.15.+q, 31.10.+z, 71.10.+x, 71.45.−d

Linear-scaling semiempirical quantum calculations for macromolecules

Tai-Sung Lee, Darrin M. York, and Weitao Yang
Department of Chemistry, Duke University, Durham, North Carolina 27708

(Received 9 April 1996; accepted 6 May 1996)

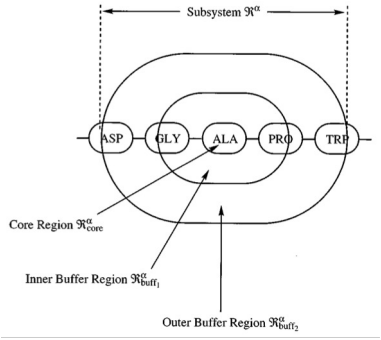
A linear-scaling method to carry out semiempirical quantum mechanical calculations for large systems has been developed based on the density matrix version of the divide-and-conquer approach. The method has been tested and demonstrated to be accurate and efficient. With this implementation, semiempirical quantum mechanical calculations are made possible for large molecules over 9000 atoms on a typical workstation. For biological macromolecules, solvent effects are included with a dielectric continuum model. © *1996 American Institute of Physics*. [S0021-9606(96)01231-7]

Semiempirical molecular orbital calculations with linear system size scaling

Steven L. Dixon and Kenneth M. Merz, Jr.
Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 21 December 1995; accepted 1 February 1996)

Details are provided for the implementation of a density matrix divide-and-conquer approximation into the framework of molecular orbital theory on nonperiodic systems. Originally developed for density functional theory, the divide-and-conquer procedure is one of the most promising in a growing list of techniques that exhibit linear scaling with respect to the number of basis functions in the system. The key to linear scaling is the division of the electronic structure calculation into a series of calculations over a set of small, overlapping subsystems. A semiempirical molecular orbital program designed around the divide-and-conquer approach has been written and a number of tests are carried out on polyglycine structures in order to evaluate its performance. For the systems examined, linear scaling is indeed observed, and the accuracy of the calculations can be controlled quite readily by the manner in which the system is divided into its component subsystems. For very large structures, the expense associated with the computation of two-center interactions will ultimately dominate the calculation, and quadratic scaling will become apparent. Techniques to linearize this aspect of the calculation are investigated and discussed. © *1996 American Institute of Physics*. [S0021-9606(96)03617-8]



Fast, accurate semiempirical molecular orbital calculations for macromolecules

Steven L. Dixon and Kenneth M. Merz, Jr.
Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 28 January 1997; accepted 9 April 1997)

A detailed review of the semiempirical divide-and-conquer (D&C) method is given, including a new approach to subsetting, which involves dual buffer regions. Comparisons are drawn between this method and other semiempirical macromolecular schemes. D&C calculations are carried out using a basic 32 Mbyte memory workstation on a variety of peptide systems, including proteins containing up to 1960 atoms. Aspects of storage and SCF convergence are addressed, and parallelization of the D&C algorithm is discussed. © *1997 American Institute of Physics*. [S0021-9606(97)01627-9]

What Can Quantum Computers Do?

ARTICLES
PUBLISHED ONLINE: 10 JANUARY 2010 | DOI: 10.1038/NCHEM.483
nature chemistry

Towards quantum chemistry on a quantum computer

B. P. Lanyon[1,2]*, J. D. Whitfield[4], G. G. Gillett[1,2], M. E. Goggin[1,5], M. P. Almeida[1,2], I. Kassal[4], J. D. Biamonte[4†], M. Mohseni[4†], B. J. Powell[1,3], M. Barbieri[1,2†], A. Aspuru-Guzik[4]* and A. G. White[1,2]

Exact first-principles calculations of molecular properties are currently intractable because their computational cost grows exponentially with both the number of atoms and basis set size. A solution is to move to a radically different model of computing by building a quantum computer, which is a device that uses quantum systems themselves to store and process data. Here we report the application of the latest photonic quantum computer technology to calculate properties of the smallest molecular system: the hydrogen molecule in a minimal basis. We calculate the complete energy spectrum to 20 bits of precision and discuss how the technique can be expanded to solve large-scale chemical problems that lie beyond the reach of modern supercomputers. These results represent an early practical step toward a powerful tool with a broad range of quantum-chemical applications.

LETTER
doi:10.1038/nature23879

Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets

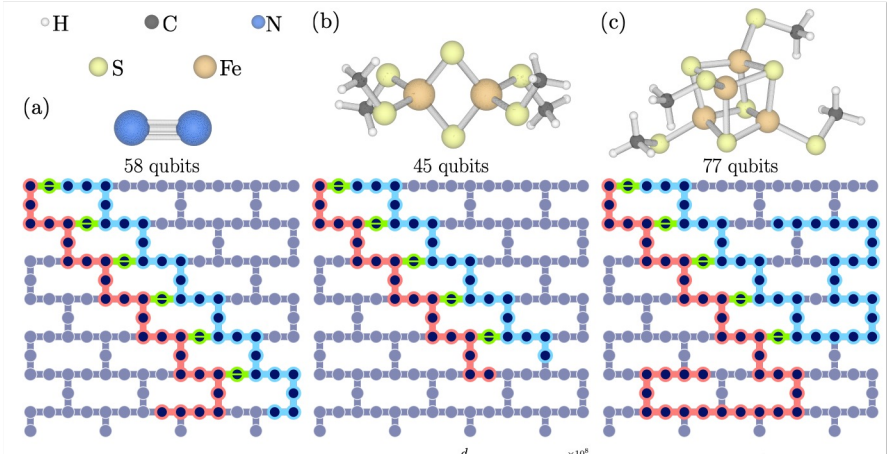
Abhinav Kandala[1]*, Antonio Mezzacapo[1]*, Kristan Temme[1], Maika Takita[1], Markus Brink[1], Jerry M. Chow[1] & Jay M. Gambetta[1]

Technological Advancements are Necessary
Sample-based Quantum Subspace Diagonalization (SQD)

Chemistry Beyond the Reach of Exact Solutions of the Schrödinger Equation on a Quantum-centric Supercomputer

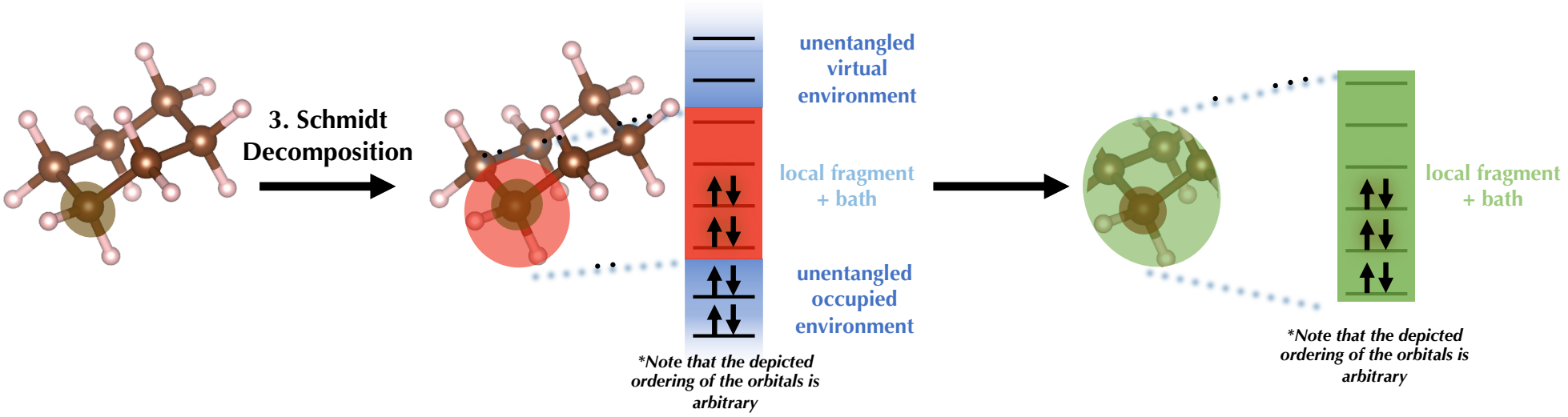
Javier Robledo-Moreno,[1,*] Mario Motta,[1,†] Holger Haas,[1] Ali Javadi-Abhari,[1] Petar Jurcevic,[1] William Kirby,[2] Simon Martiel,[3] Kunal Sharma,[1] Sandeep Sharma,[4] Tomonori Shirakawa,[5,6,7] Iskandar Sitdikov,[1] Rong-Yang Sun,[5,6,7] Kevin J. Sung,[1] Maika Takita,[1] Minh C. Tran,[2] Seiji Yunoki,[5,6,7,8] and Antonio Mezzacapo[1,‡]

[1] *IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598*
[2] *IBM Quantum, IBM Research Cambridge, Cambridge, MA 02142, USA*
[3] *IBM Quantum, IBM France Lab, Orsay, France*
[4] *Department of Chemistry, University of Colorado, Boulder, CO 80302, USA*
[5] *Computational Materials Science Research Team, RIKEN Center for Computational Science (R-CCS), Kobe, Hyogo, 650-0047, Japan*
[6] *Quantum Computational Science Research Team, RIKEN Center for Quantum Computing (RQC), Wako, Saitama, 351-0198, Japan*
[7] *RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), Wako, Saitama 351-0198, Japan*
[8] *RIKEN Center for Emergent Matter Science (CEMS), Wako, Saitama 351-0198, Japan*



https://arxiv.org/abs/2405.05068;
https://www.science.org/doi/10.1126/sciadv.adu9991

EWF (Embedded Wavefunction Theory) Workflow



1. Hartree Fock Calculation on full molecule
2. Choose fragment orbitals

Vayesta + pyscf

4. Select the bath orbitals to form the cluster (fragment+bath) – η - value

5. Rotate cluster orbitals into occupied/unoccupied blocks

6. Expand the cluster with fragment-specific natural orbitals – "entangle" cluster with full environment

7. High Level Quantum Chemistry Solver

*SQD, CI, or CCSD

**Single step to assemble the system

8. Get the high level RDMs to form the total energy

$$E_{total} = E_{HF} + \sum_{x}^{N_F} E_{corr}^{x}$$

https://pyscf.org
https://qiskit.github.io/qiskit-addon-sqd
https://github.com/sandbox-quantum/Tangelo

Protein Conformational Ensembles

Trp-Cage

Protein Ensembles and disease

1) Intrinsically disordered proteins (IDPs)
   - Alzheimer's, Parkinson's, Huntington's and ALS
   - Involved in cancer pathways
   - Cardiovascular disease and diabetes
2) Allosteric pockets
   - Druggable sites beyond the active site
   - Hard to find and exploit
3) Structure-based drug discovery

Trp-Cage

1) Designed folded min-protein (Neidigh et al, 2002)
2) Fast folder
3) Folding free energy is ~2.1kcal/mol (in aqueous solution at 300K)
4) One of the first proteins to be folded using MD simulations

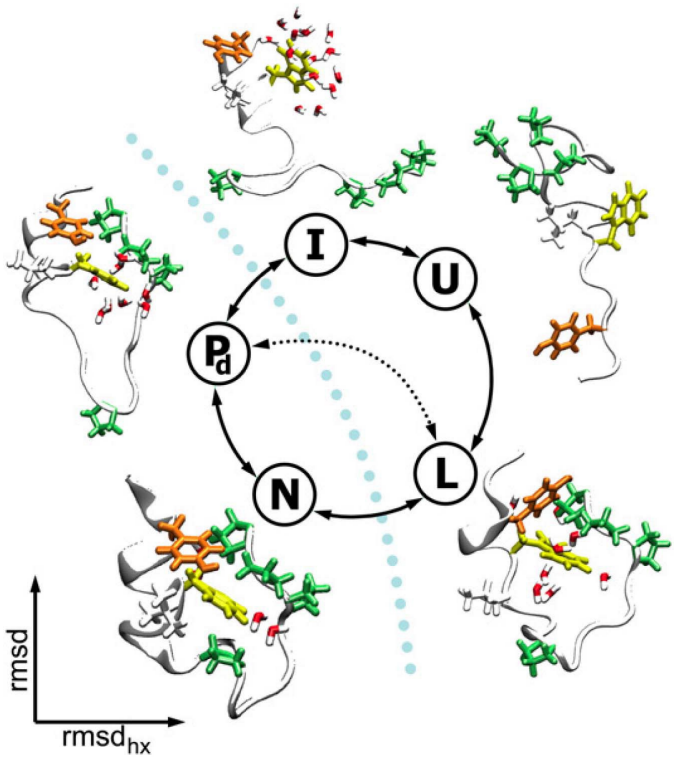
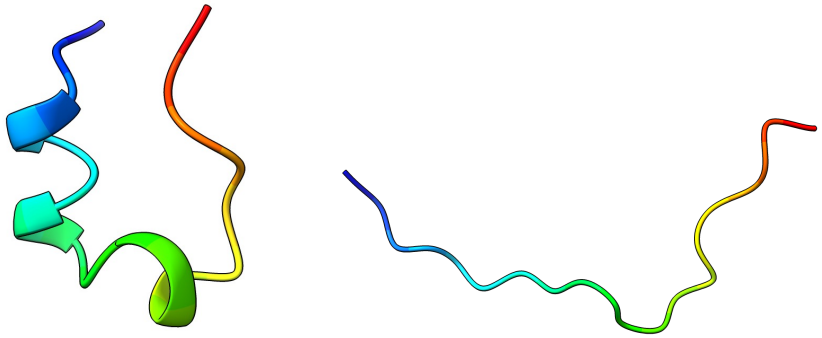


Image see: https://www.cell.com/fulltext/S0006-3495(08)78564-3

Hybrid Quantum-Centric Energy Evaluation

Trp-cage protein



Folded Conformer **Unfolded Conformer**

Quantum mechanical evaluation of relative energy (ΔE) of two Trp-cage protein conformers using embedded wavefunction (EWF) method on IBM Heron R2 hardware.

- Execute quantum circuits of fragments on IBM Heron R2.
- Perform energy calculations for each conformer.
- Combine quantum-classical results to compute ΔE between two conformers.

Total Number of Quantum Measurements on IBM Heron R2 hardware for Two Conformers using >300 circuits

300 Million shots

Distinct Quantum Circuits Executed for Two Conformers

> 300 Circuits

Number of Qubits used across 40+ Circuits

60+ qubits/circuit

Total Number of Nodes and CPUs used simultaneously

>150 Nodes and 7200 CPUs 7K Node Hours

IBM Quantum

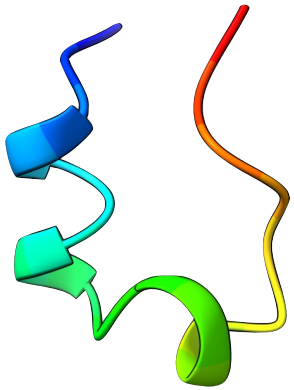
IBM Quantum

Harnessing GPU-QPU-CPU in a Single Workflow

GPU+QPU+CPU

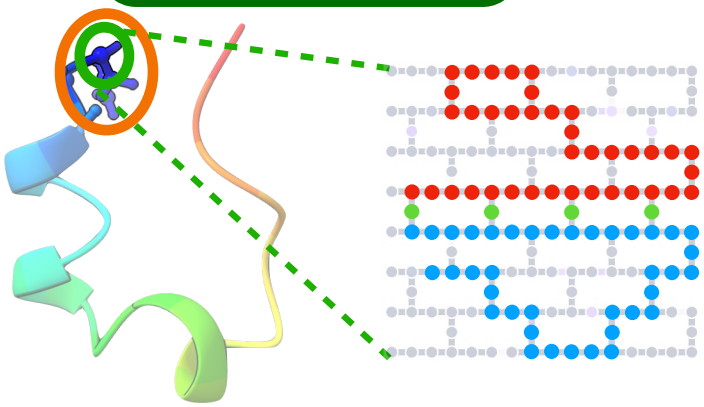
GPU

Simulation of Initial Mean Field Object on full system



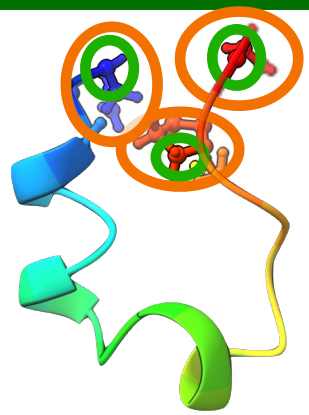
QPU

Collecting candidate electron configurations from each LUCJ circuit

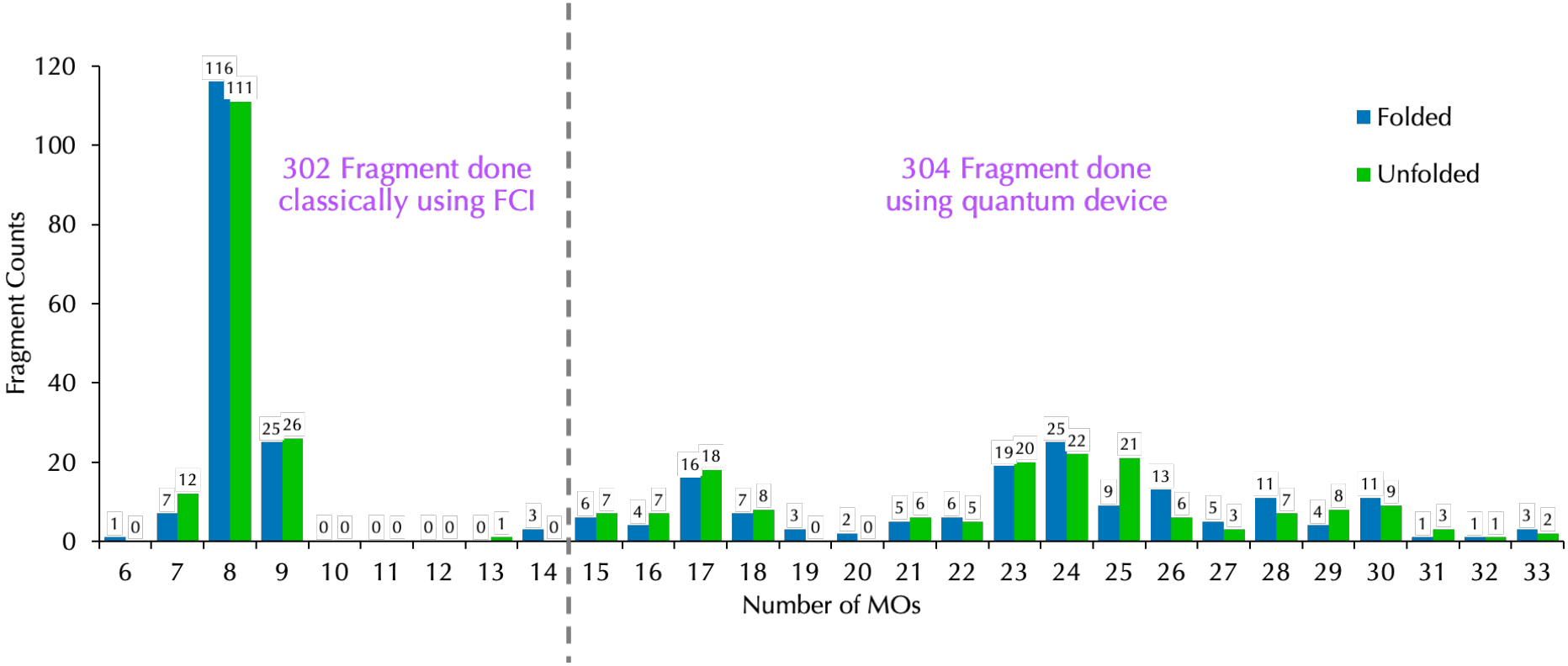


CPU

Post-processing of the Data from Quantum Computer

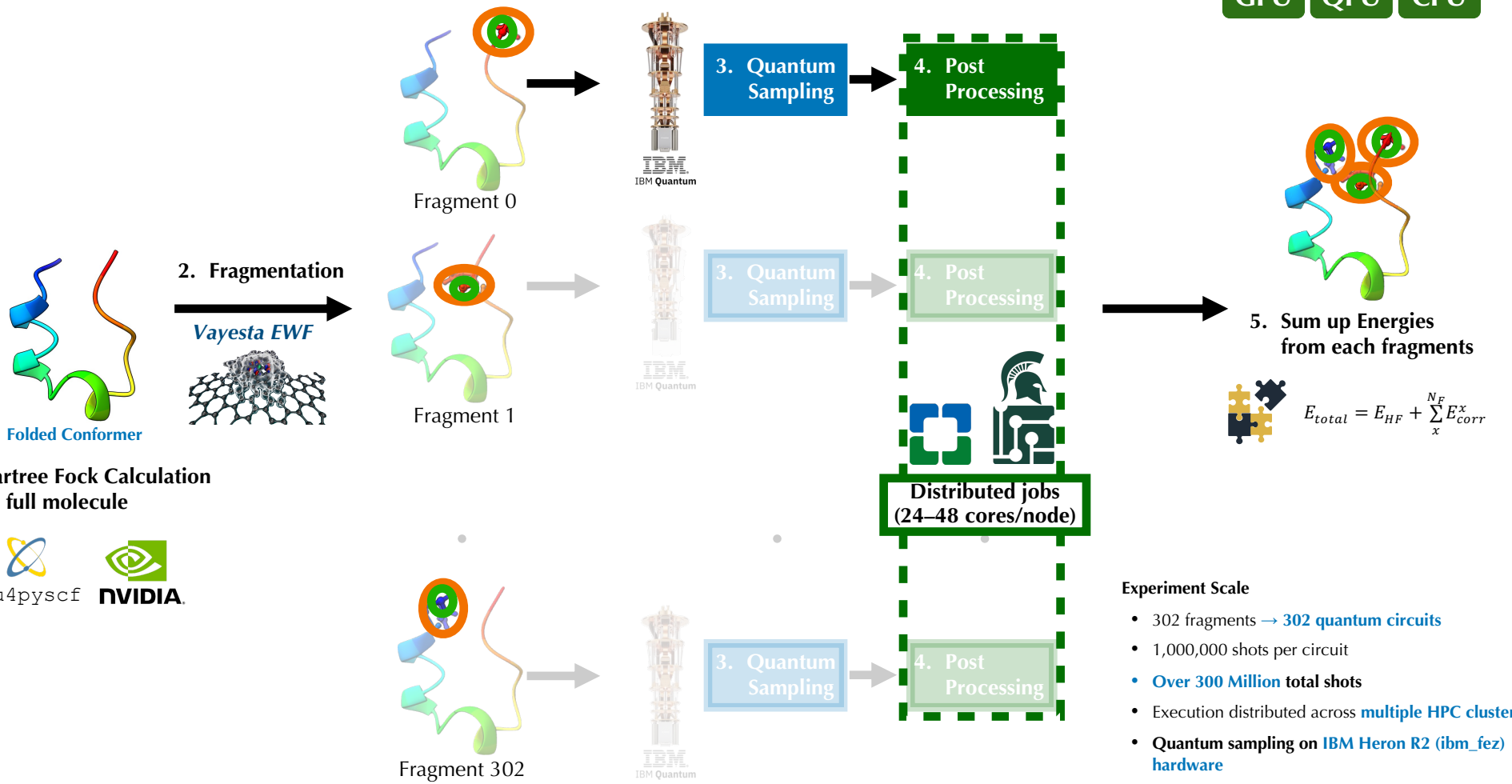


Fragment Distribution Across Different Sizes of MOs



Distributed Hybrid Quantum-centric Workflow

GPU QPU CPU



1. **Hartree Fock Calculation on full molecule**

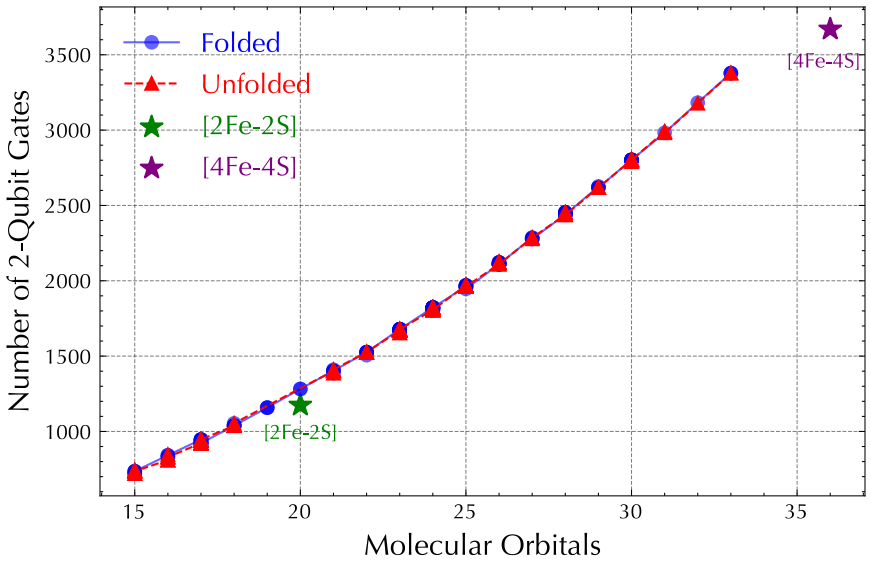
gpu4pyscf NVIDIA

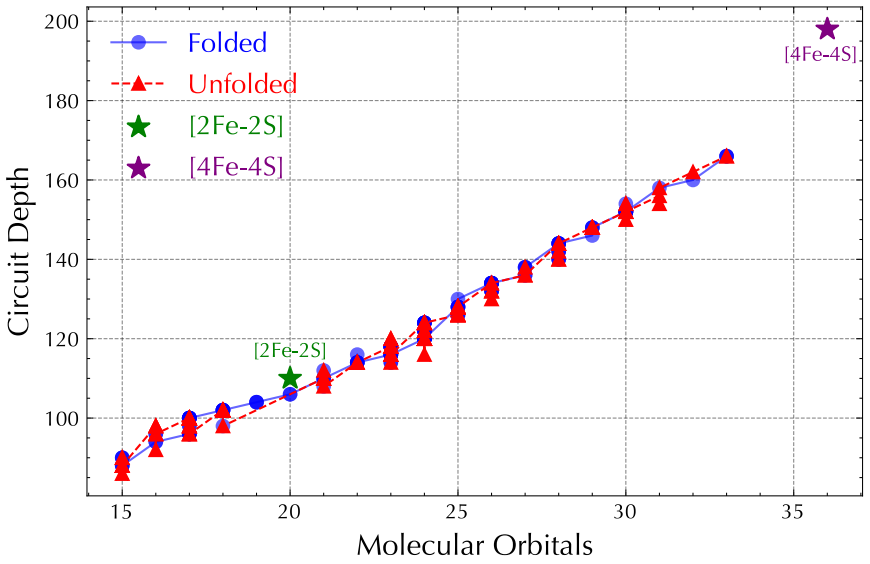
Experiment Scale

- 302 fragments → **302 quantum circuits**
- 1,000,000 shots per circuit
- **Over 300 Million total shots**
- Execution distributed across **multiple HPC clusters**
- **Quantum sampling on IBM Heron R2 (ibm_fez) hardware**

Quantum Computing Complexity of the Fragments

GPU QPU CPU





- The complexity of quantum circuits increases with the size of fragments
- The quantum circuit complexity in the largest fragment approaching the complexity of most advanced SQD simulation to date performed in 4Fe-4S cluster

Relative Energies of Conformers (ΔE)

STO-3G basis set

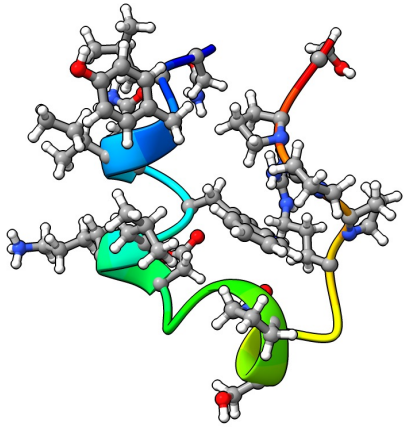
Reference method RI-MP2	EWF-MP2
-63.42	*-59.61* *(ΔΔE = 3.8)*

Reference method DLPNO-CCSD	EWF-CCSD
-52.05	*-47.70* *(ΔΔE = 4.4)*

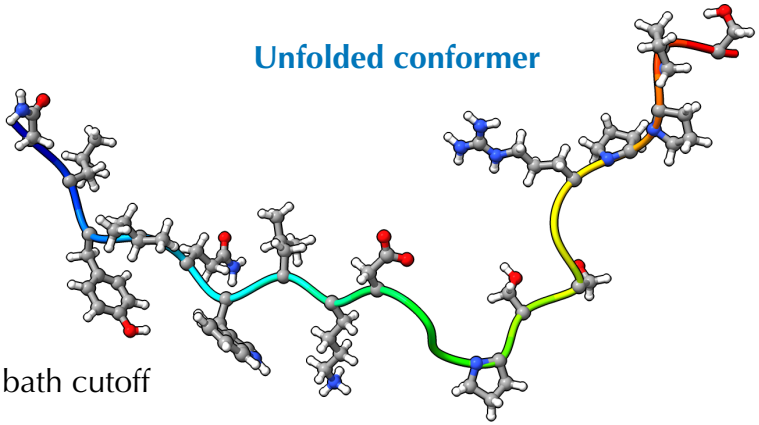
Reference method DLPNO-CCSD	EWF-(ext-SQD-FCI)*
-52.05	*-55.43* *(ΔΔE = 3.4)*

*The EWF-(ext-SQD-FCI) energy uses FCI for MOs<15 and SQD for MOs>=15; 1e-5 bath cutoff

Folded conformer

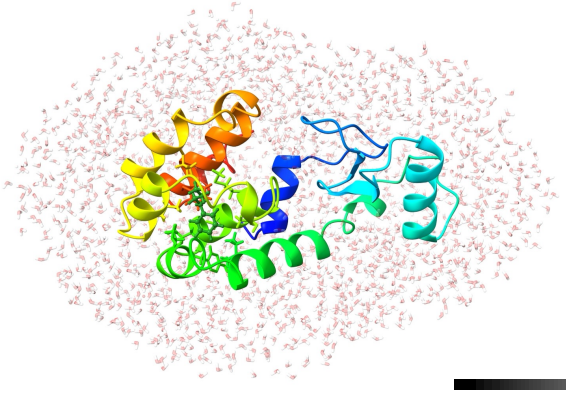


Unfolded conformer

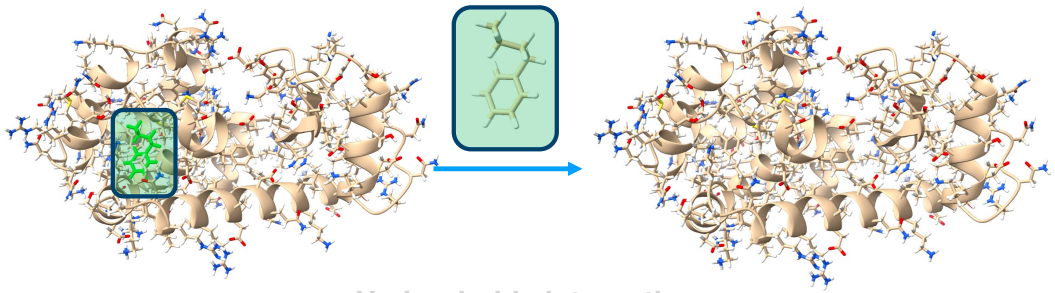


Protein Ligand Interaction

Solvated T4-Lysozyme

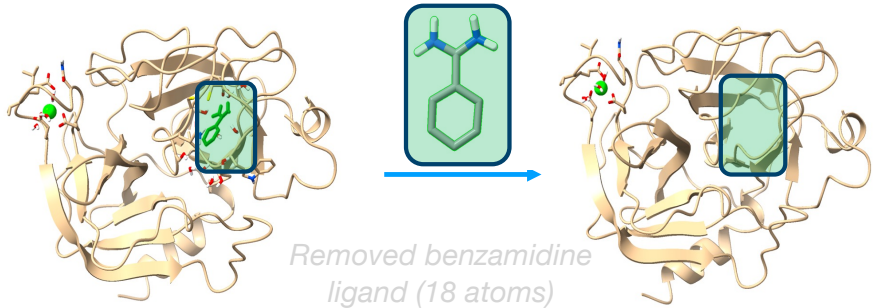


Removed n-butylbenzene ligand (24 atoms)



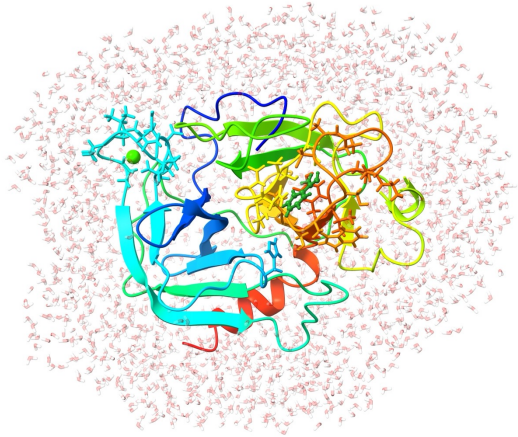
Hydrophobic interaction

Hydrophilic interaction

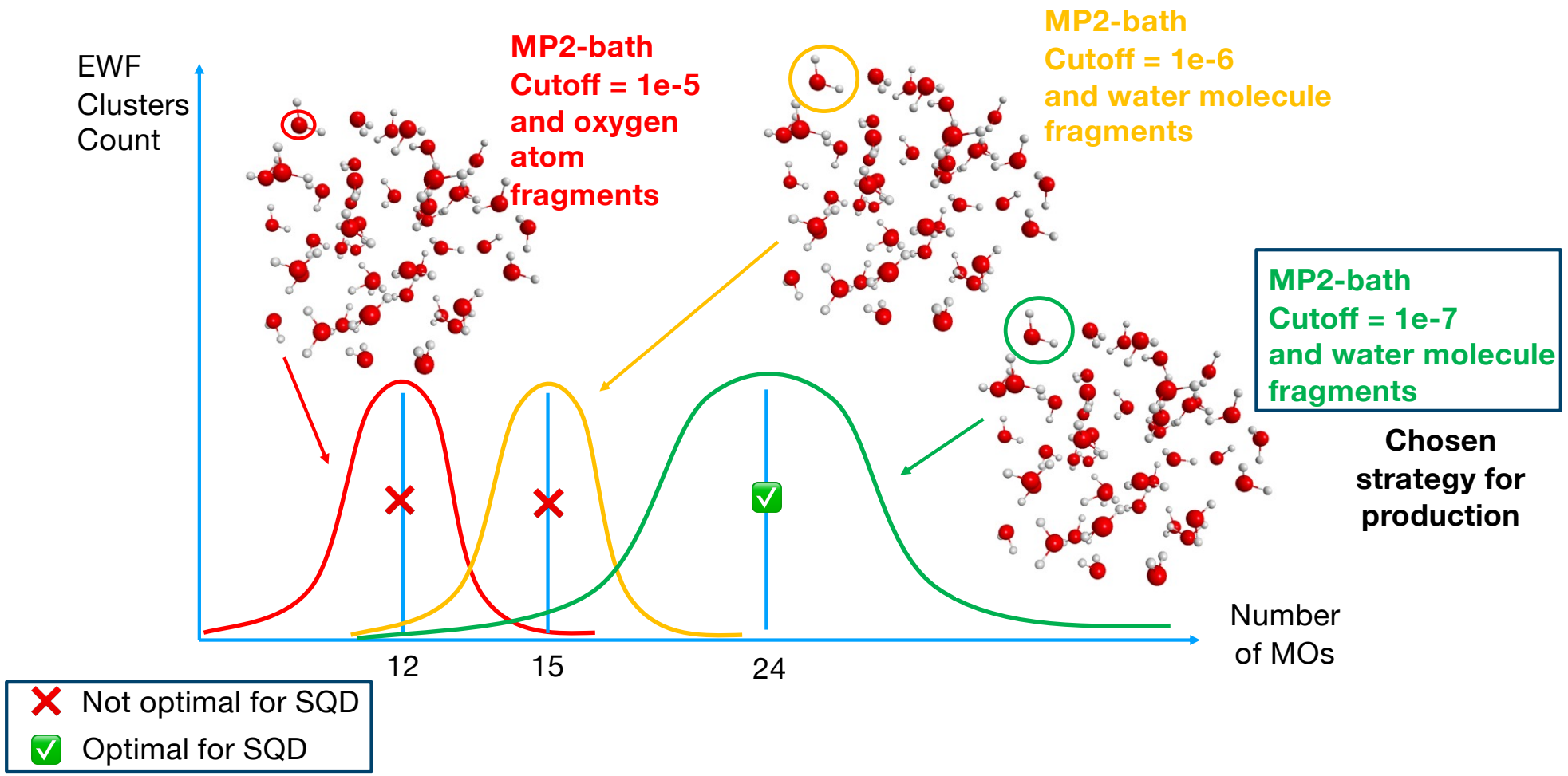


Removed benzamidine ligand (18 atoms)

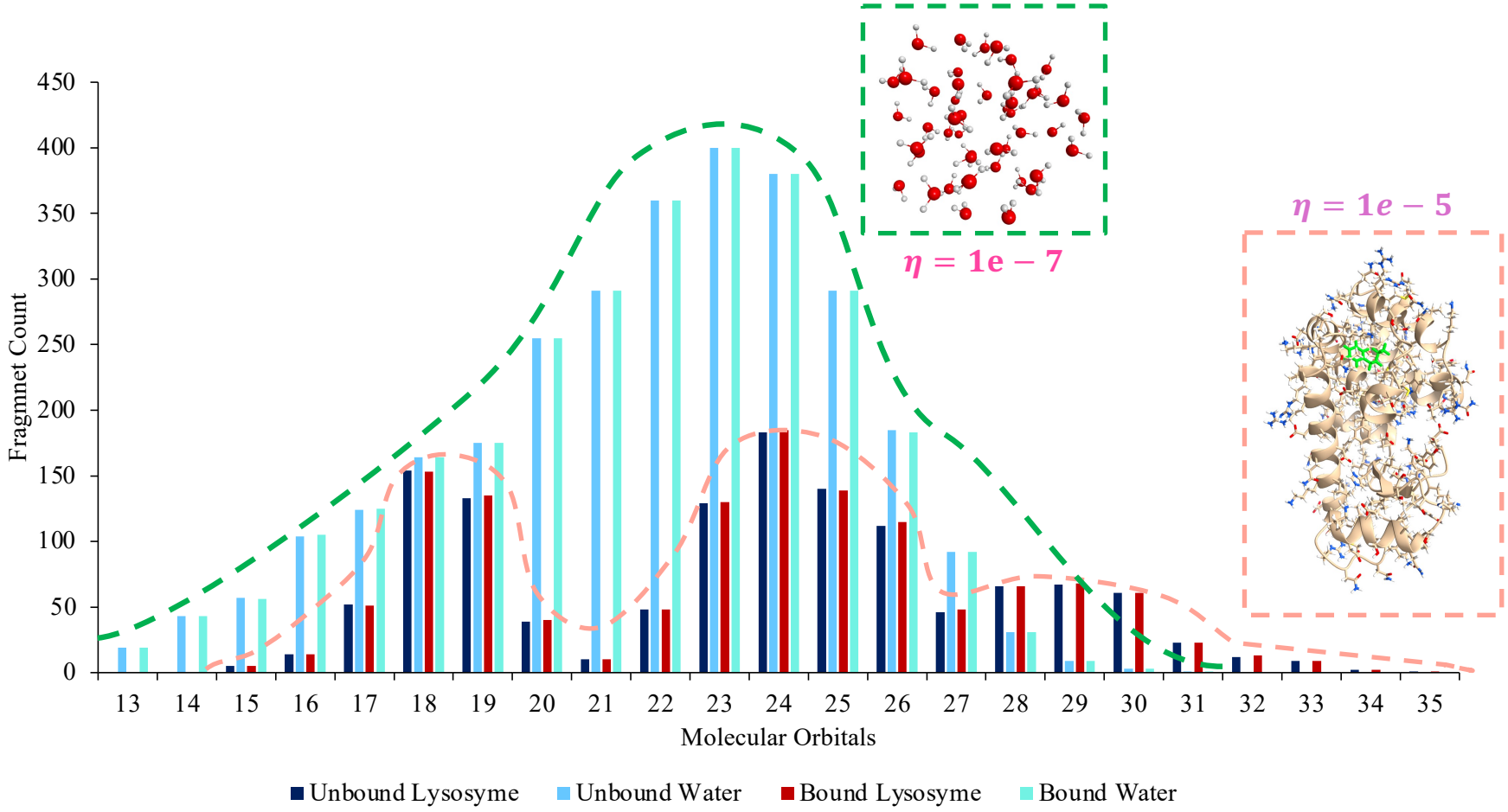
Solvated Trypsin



Strategies for Water EWF Clusters



Fragmentation MO distribution from Production run: T4-Lysosyme



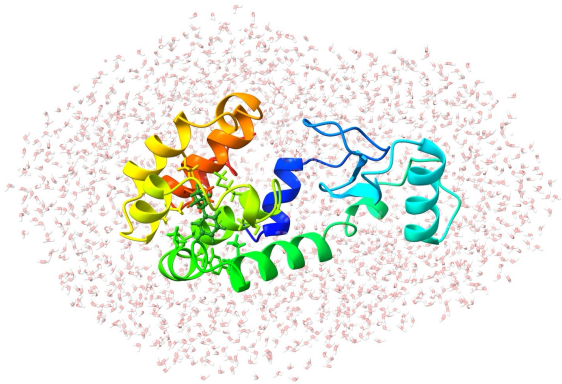
Slide credits: Danil Kaliakin

T4-Lysosyme/N-butylbenzene Results

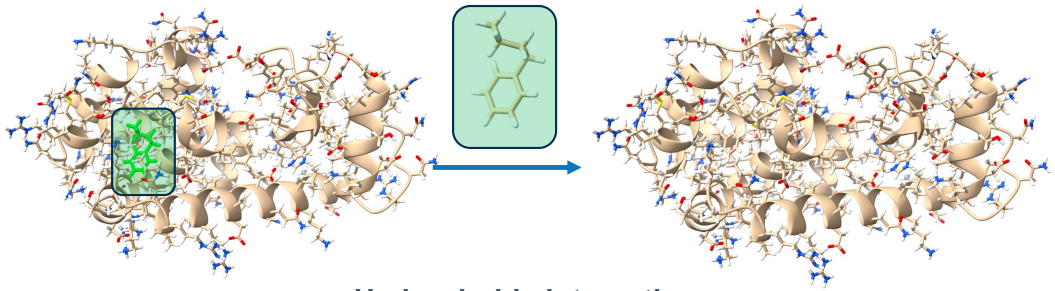
	HF	EWF-MP2	EWF-CCSD	EWF-(TrimSQD-FCI)*
Bound	-288932.37217	-289098.83683	-289139.28385	-289141.60903
Unbound	-288550.21447	-288716.05224	-288756.38651	-288758.70867
Ligand	-382.17691	-382.79189	-382.91398	-382.91518
E_Bind (Ha)	0.01921	0.00730	0.01665	0.01483
E_Bind (kcal/mol)	12.06	4.58	10.45	9.30

**All protein hydrogen fragments are treated with FCI. For other fragments with 13–26 MOs, Extended SQD is used (D=3,000, $\varepsilon=1\times10^{-6}$); for larger fragments with 27–35 MOs, Trimmed SQD is applied (D=10,000, $\varepsilon=5\times10^{-6}$).*

Solvated T4-Lysozyme

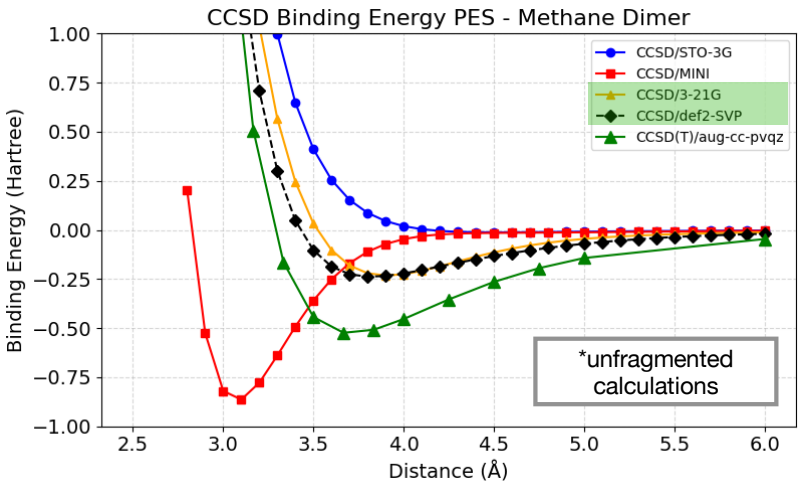


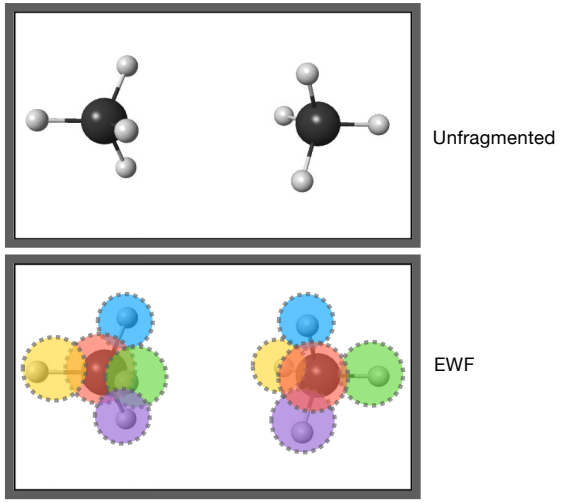
Removed n-butylbenzene ligand (24 atoms)

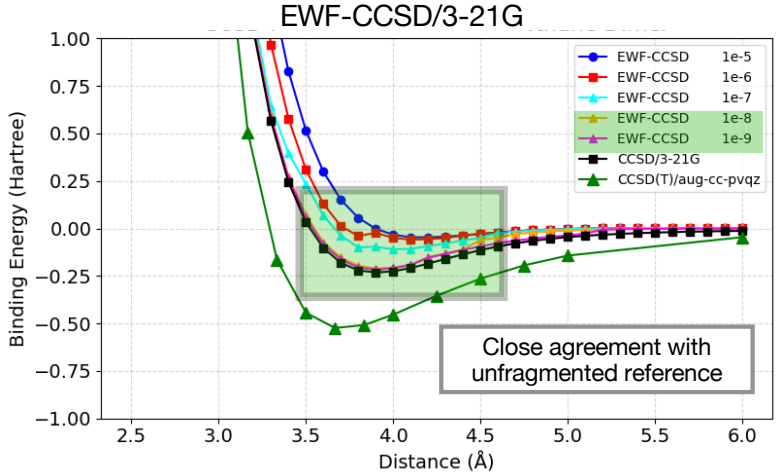


Hydrophobic interaction

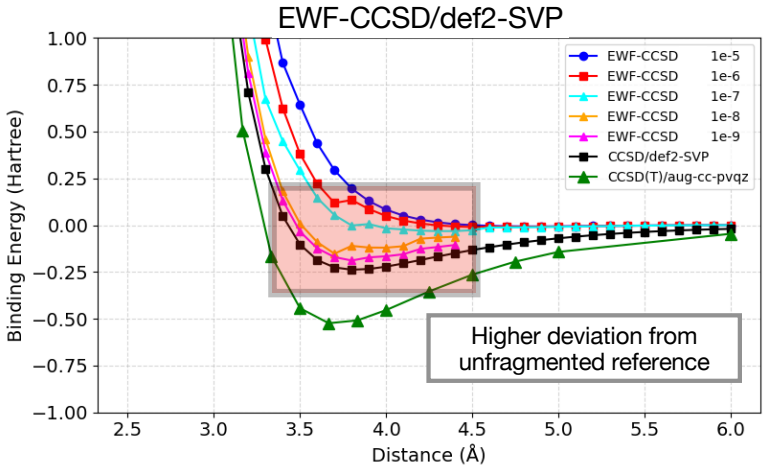
Choice of Basis Set and EWF Threshold







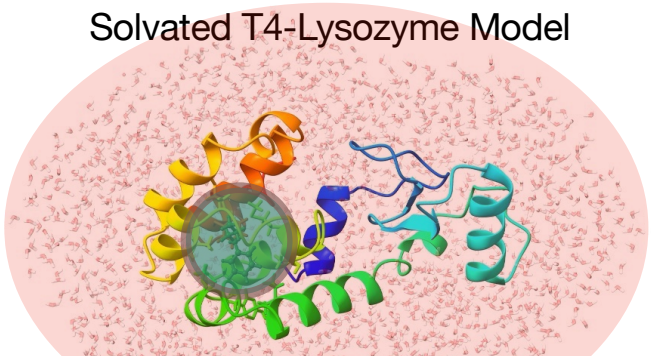
Effect of EWF bath threshold



Effect of EWF bath threshold

Mixed Basis Set Simulations

Solvated T4-Lysozyme Model



Residues within 3 Å radius are treated with 3-21G

3-21G
η = 1e-8

sto-3g
η = 1e-5

η is EWF bath threshold

	STO-3G only		3-21G and STO-3G	
	EWF-CCSD	EWF–TrimSQD	EWF-CCSD	EWF–TrimSQD
Bound	-289139.284	-289141.609	-289179.536	-289182.051
Unbound	-288756.387	-288758.709	-288793.907	-288796.420
Ligand	-382.914	-382.915	-385.601	-385.598
E_Bind **(Ha)**	0.017	0.015	-0.028	-0.033
E_Bind **(kcal/mol)**	**10.45**	**9.30**	**-17.31**	**-21.07**

Mixed Basis Set Simulations

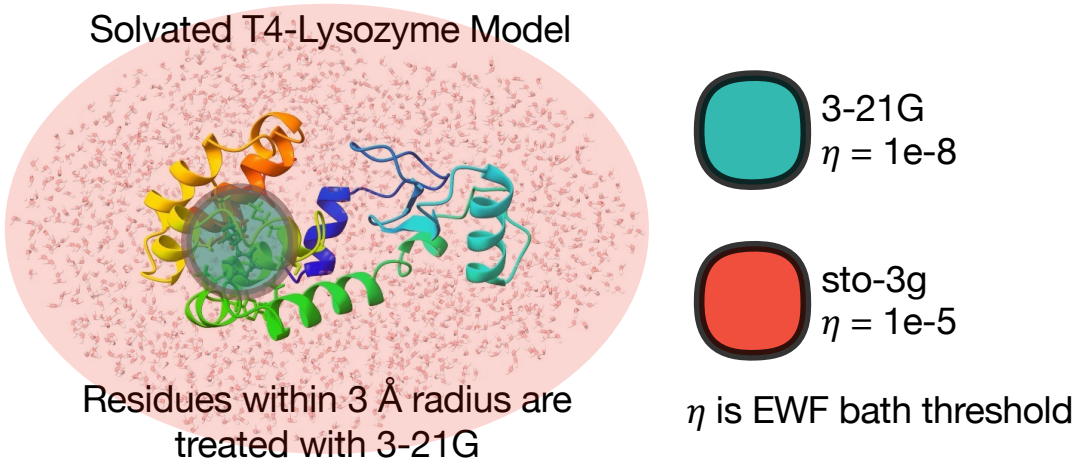


TABLE VIII

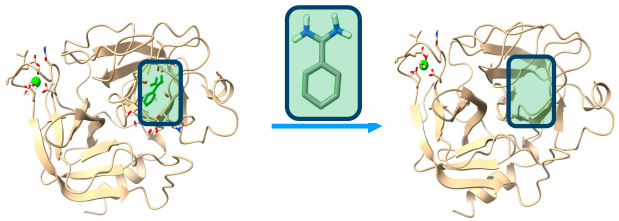
TOTAL AND BINDING ENERGIES, DEFINED AS $E_{\text{binding}} = E_{\text{bound}} - E_{\text{unbound}} - E_{\text{ligand}}$, OF THE BENCHMARK MOLECULAR SYSTEMS.

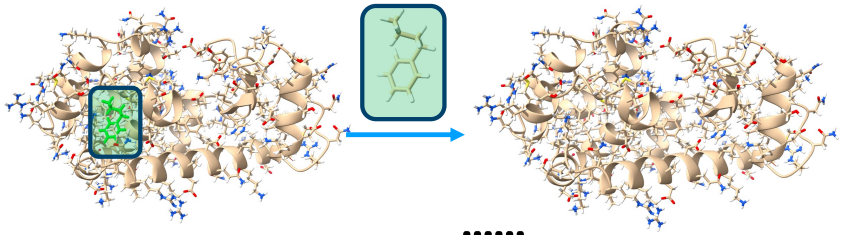
System	**Method**	$\mathbf{E}_{\text{bound}}$ (Ha)	$\mathbf{E}_{\text{unbound}}$ (Ha)	$\mathbf{E}_{\text{ligand}}$ (Ha)	$\mathbf{E}_{\text{binding}}$ (Ha)	$\mathbf{E}_{\text{binding}}$ (kcal/mol)
Trypsin, benzamidine (STO-3G)	EWF–CCSD	$-319,413.0227$	$-319,038.0806$	-374.9875	0.0454	28.46
	EWF–TrimSQD	$-319,415.8966$	$-319,040.9552$	-374.9986	0.0572	35.89
T4-Lysozyme, n-butyl-benzene (STO-3G)	EWF–CCSD	$-289,139.2838$	$-288,756.3865$	-382.9139	0.0166	10.45
	EWF–TrimSQD	$-289,141.6090$	$-288,758.7087$	-382.9152	0.0148	9.30
Trypsin, benzamidine (STO-3G, electrostatic correction)	EWF–CCSD	$-319,413.0227$	$-319,037.8220$	-374.9926	-0.2081	-130.59
	EWF–TrimSQD	$-319,415.8966$	$-319,040.6832$	-375.0043	-0.2092	-131.32
T4-Lysozyme, n-butyl-benzene (mixed 3-21G and STO-3G)	EWF–CCSD	$-289,179.5355$	$-288,793.9070$	-385.6009	-0.0275	-17.31
	EWF–TrimSQD	$-289,182.0538$	$-288,796.4222$	-385.5980	-0.0336	-21.07

Scope Of The Work

Total number of quantum measurements on IBM Heron R2 hardware

Over 1.3 billion shots





Distinct quantum circuits executed

> 9200 circuits

Number of circuits with 60+ qubits

1000+ circuits

Total number of CPUs used simultaneously

>158000 CPUs

Expansion of methodology to simulations of binding energy in two complex protein systems corresponding to hydrophobic and hydrophilic interactions between protein and ligand. Associated simulations are significantly more complex than current state of the art both in classical and quantum complexity.

2026 ACM Gordon Bell Prize Finalist

Cleveland Clinic · RIKEN · IBM — 12,635-atom quantum-enabled protein simulation

- **First IBM GB finalist in nearly a decade**
- **First-ever quantum computers in GB**
- **Largest quantum-enabled bio-molecular simulation**

QUANTUM-CENTRIC SUPERCOMPUTING RESOURCES

~190 hours QPU time
IBM Cleveland & Kobe
Heron (Cleveland) Heron (Kobe)

4M node-hours
RIKEN Fugaku (CPU)

~20K GPU node-hours
Miyabi-G & ROQUO

12,635 atoms simulated
Largest quantum bio-sim

END-TO-END WORKFLOW INNOVATION

BEFORE ~3 months

1. Generate fragments at Cleveland Clinic
2. Sample on quantum computers (Cleveland & Kobe)
3. Manually copy data to RIKEN
4. Diagonalize on Fugaku / Miyabi-G
5. Transfer results back to Cleveland for final energies

Many people, ad-hoc coordination across 3 organizations

AFTER ~65 hours ~33x faster

- Fully automated orchestration across QPUs and HPC
- Quantum sampling and classical diagonalization run concurrently
- Runs on RIKEN's ROQUO GPU system + IBM Quantum
- Reproducible, production-grade workflow

Goal: democratize this workflow so 10K–100K atom simulations are routine

COLLABORATION & IMPACT

Cleveland Clinic
Computational chemistry
Domain science & proteins

RIKEN
HPC infrastructure
Fugaku · ROQUO · expertise

IBM
Quantum hardware & software
Heron processors · Qiskit

SCIENTIFIC IMPACT

Better accuracy
Binding energies now reflect expected behavior in benchmark systems

Scalable science
Study more proteins, more conditions, more complex molecules

Demonstration → Discovery
From proof-of-concept to production science

Final presentation at SC26 · November 2026 · Chicago, IL · One of six Gordon Bell finalists

arXiv > quant-ph > arXiv:2512.17130

Quantum Physics

[Submitted on 18 Dec 2025]

Molecular Quantum Computations on a Protein

Akhil Shajan, Danil Kaliakin, Fangchun Liang, Thaddeus Pellegrini, Hakan Doga, Subhamoy Bhowmik, Susanta Das, Antonio Mezzacapo, Mario Motta, Kenneth M. Merz Jr

This work presents the implementation of a fragment-based, quantum-centric supercomputing workflow for computing molecular electronic structure using quantum hardware. The workflow is applied to predict the relative energies of two conformers of the 300-atom Trp-cage miniprotein. The methodology employs wave function-based embedding (EWF) as the underlying fragmentation framework, in which all atoms in the system are explicitly included in the CI treatment. CI calculations for individual fragments are performed using either sample-based quantum diagonalization (SQD) for challenging fragments or full configuration interaction (FCI) for trivial fragments. To assess the accuracy of SQD for fragment CI calculations, EWF-(FCI,SQD) results are compared against EWF-MP2 and EWF-CCSD benchmarks. Overall, the results demonstrate that large-scale electronic configuration interaction (CI) simulations of protein systems containing hundreds or even thousands of atoms can be realized through the combined use of quantum and classical computing resources.

Subjects: **Quantum Physics (quant-ph)**
Cite as: arXiv:2512.17130 **[quant-ph]**
(or arXiv:2512.17130v1 **[quant-ph]** for this version)
https://doi.org/10.48550/arXiv.2512.17130

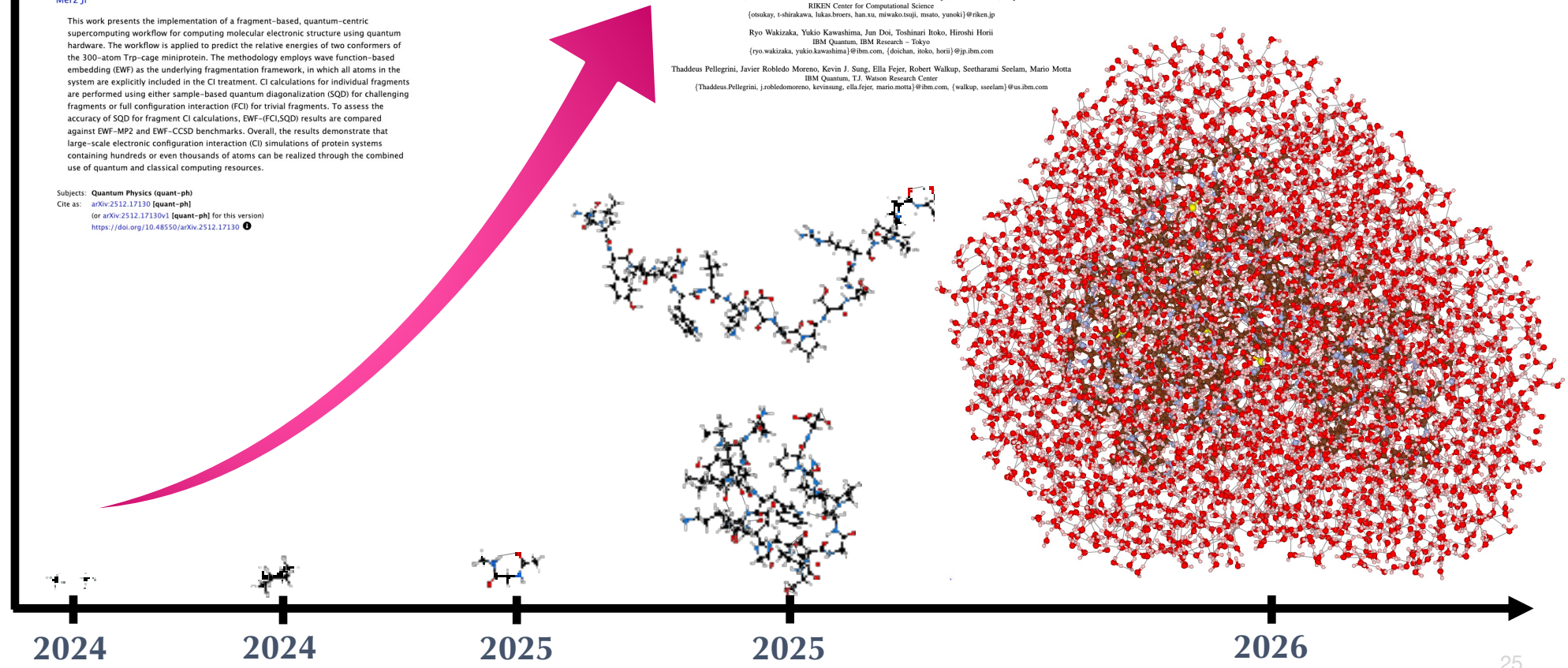
Crossing the 12,000-atom barrier with heterogeneous quantum-classical supercomputing: quantum chemistry of protein-ligand complexes

Kenneth M. Merz, Jr., Akhil Shajan, Danil Kaliakin, Fangchun Liang
Cleveland Clinic Foundation
{merzk, shajana, kaliakd, liangf}@ccf.org

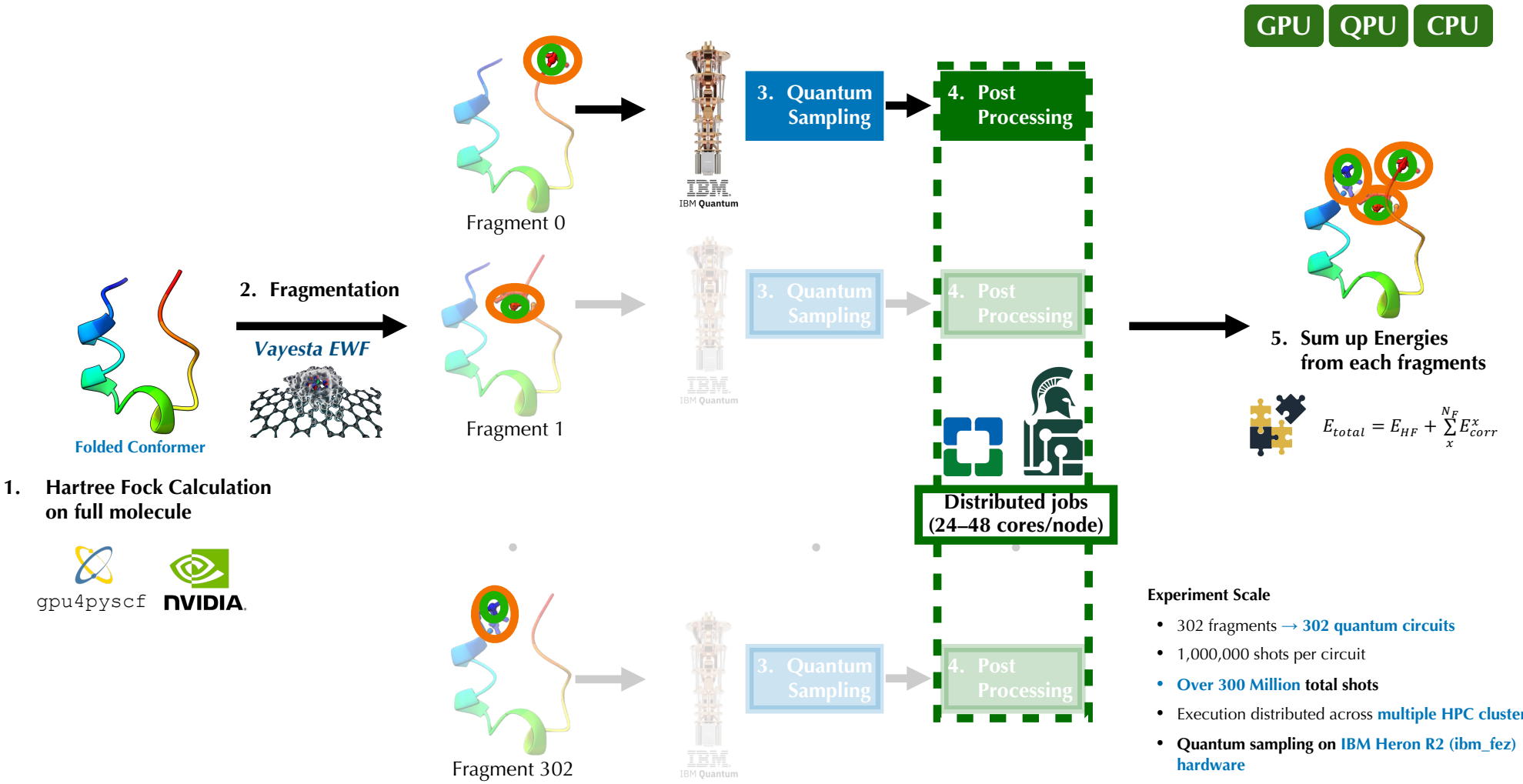
Yuichi Otsuka, Tomonori Shirakawa, Lukas Broers, Han Xu, Miwako Tsuji, Mitsuhisa Sato, Seiji Yunoki
RIKEN Center for Computational Science
{otsukay, t-shirakawa, lukas.broers, han.xu, miwako.tsuji, msato, yunoki}@riken.jp

Ryo Wakizaka, Yukio Kawashima, Jun Doi, Toshinari Itoko, Hiroshi Horii
IBM Quantum, IBM Research – Tokyo
{ryo.wakizaka, yukio.kawashima}@ibm.com, {doichan, itoko, horii}@jp.ibm.com

Thaddeus Pellegrini, Javier Robledo Moreno, Kevin J. Sung, Ella Fejer, Robert Walkup, Seetharami Seelam, Mario Motta
IBM Quantum, T.J. Watson Research Center
{Thaddeus.Pellegrini, j.robledomoreno, kevinsung, ella.fejer, mario.motta}@ibm.com, {walkup, sseelam}@us.ibm.com



Flexible Quantum-centric Supercomputing Workflow



Experiment Scale

- 302 fragments → **302 quantum circuits**
- 1,000,000 shots per circuit
- **Over 300 Million total shots**
- Execution distributed across **multiple HPC clusters**
- **Quantum sampling on IBM Heron R2 (ibm_fez) hardware**

Summary

- Largest GPU/QPU/CPU QCSC calculation to date.
- Interfacing QC with statistical mechsnics
- Workflow defined – clear pathways for improvement both classical and quantum!
- Access to quantum hardware is key!
  - USA and Europe need to increase quantum access.
- Need to rethink how we approach problems.
  - Out with the old algorithms and in with the quantum.
  - We need more circuit optimization focused on applications.
- Open-source ecosystem was key to make fast progress!
  - Unified software stack is essential (IBM/CCF/Riken)


Is QCSC at the end of the Rainbow?