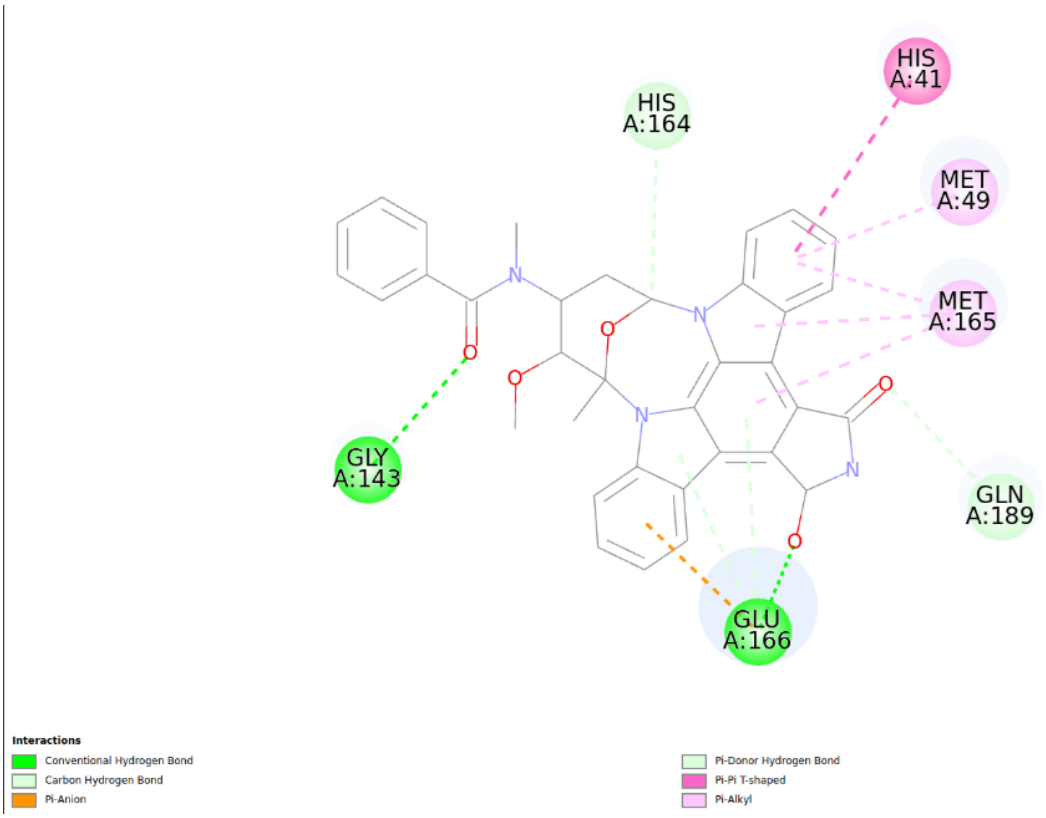
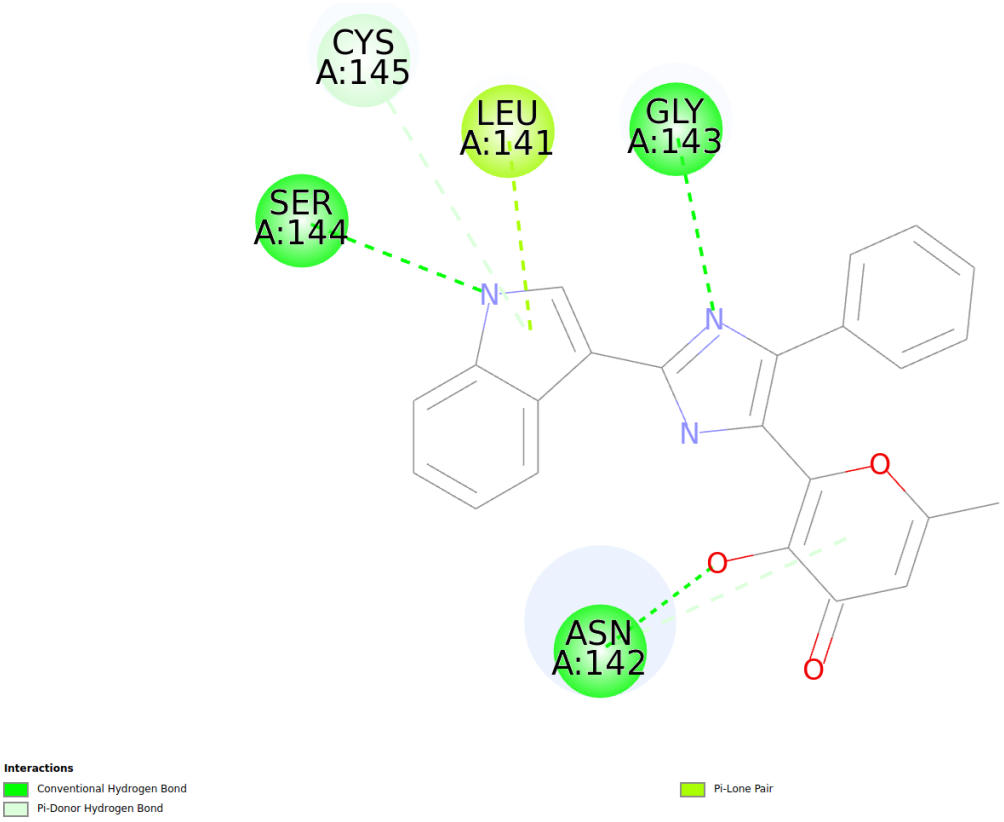
Introduction

I introduce here a new docking scheme that addresses a major source of error in conventional docking—namely, the estimation of low-energy conformations and conformational strain energies. As initial examples, I have recently published results on several FDA-approved drugs and natural products that could serve as potent inhibitors against COVID-19. The figures below illustrate one example from each category.

The visualizations highlight different intermolecular interactions between the S stereoisomer of 3-hydroxy Midostaurin and the 6Y2G target protein, using poses obtained from the QFVinardo hit lists.



Visualizations of different intermolecular interactions in between CNP0415554 (from COCONUT database) and the 6LU7 target protein using the pose from the hit lists of QFVinardo.

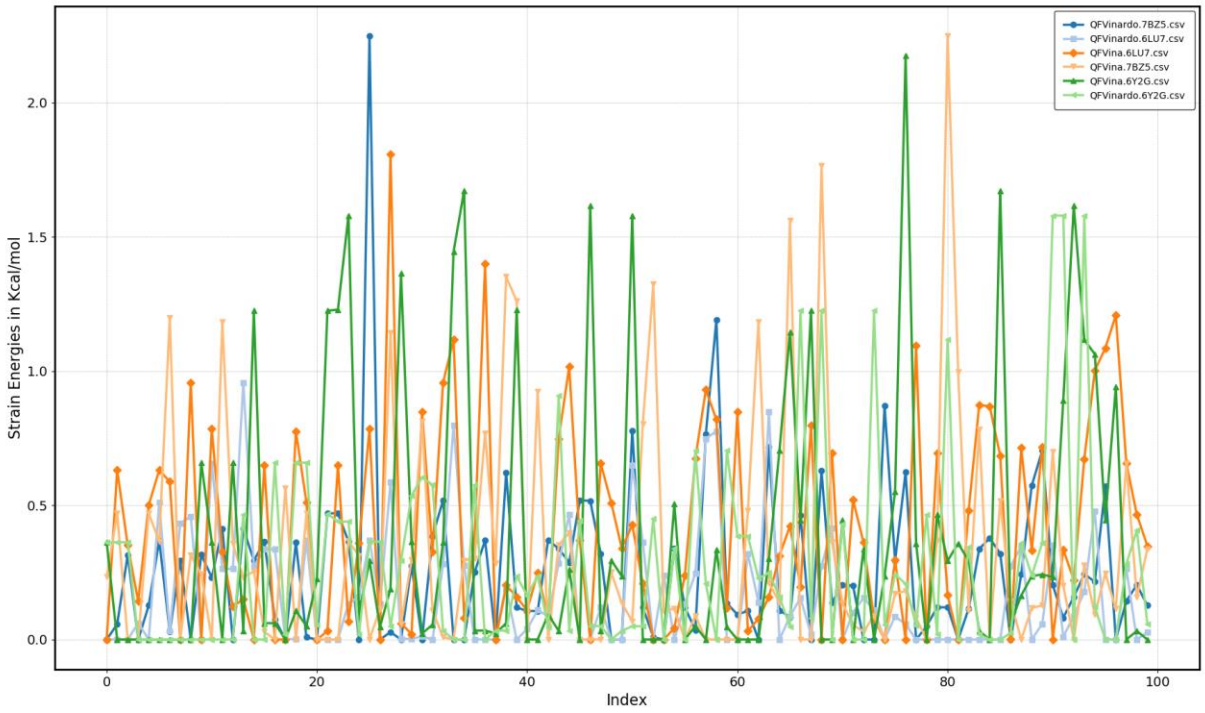


This technique employs an accurate precalculated conformational database, which is developed through the following steps:

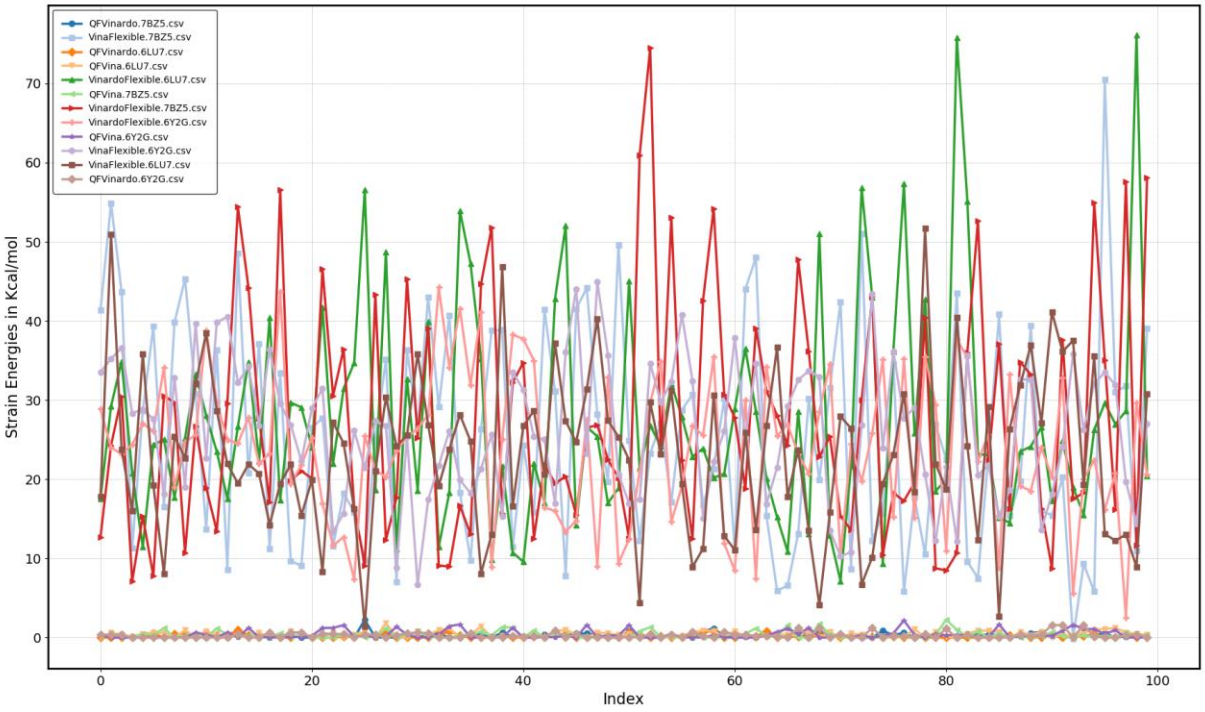
1. A robust conformational space search using three distinct force fields, followed by
2. Quantum mechanical (QM) geometry optimizations and reranking, and finally
3. DFT-D4 energy calculations using the highly accurate rev-SCAN functional and a large def2-TZVP basis set. The results are reranked, and the DFT-D4 conformational energies are attached to the final SDF file.

I will present a detailed analysis of the conformational strain energies of the docked ligand, comparing traditional Vina and Vinardo scoring functions with the QFVina and QFVinardo scoring functions. The latter replace the internal ligand energies with DFT-D4-based relative conformational energies. The following three figures illustrate these differences clearly.

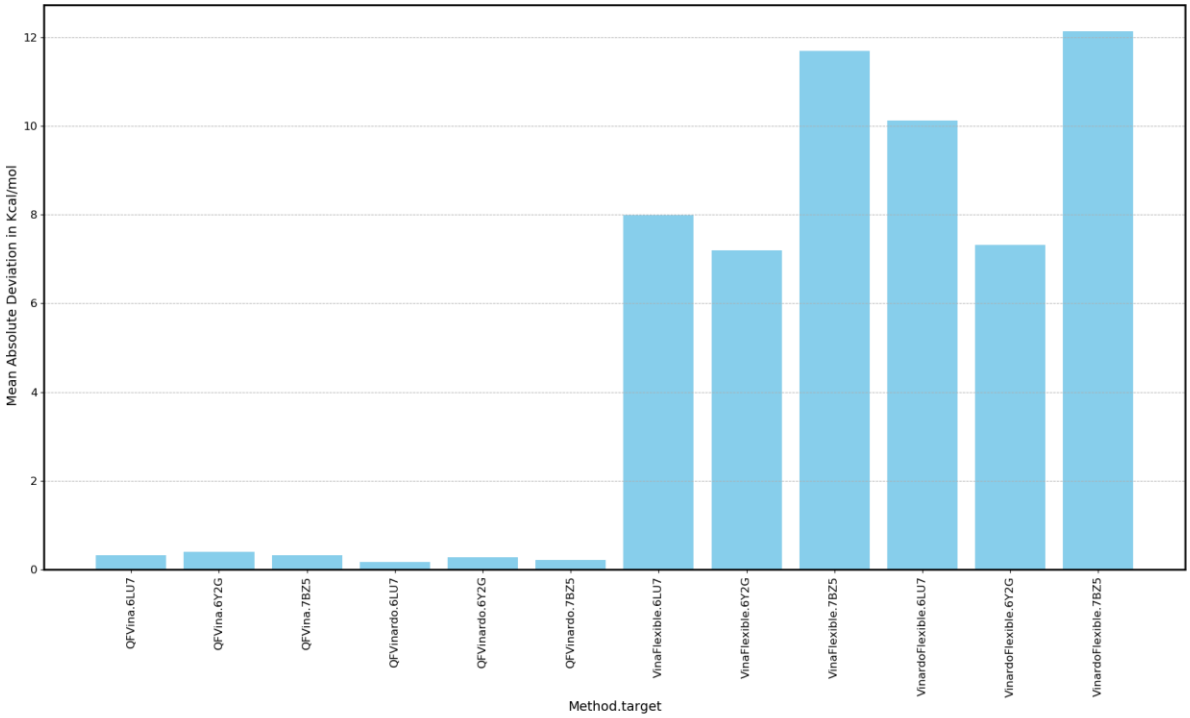
Conformational strain energies of the top100 hits in QFVina and QFVinardo dockings.



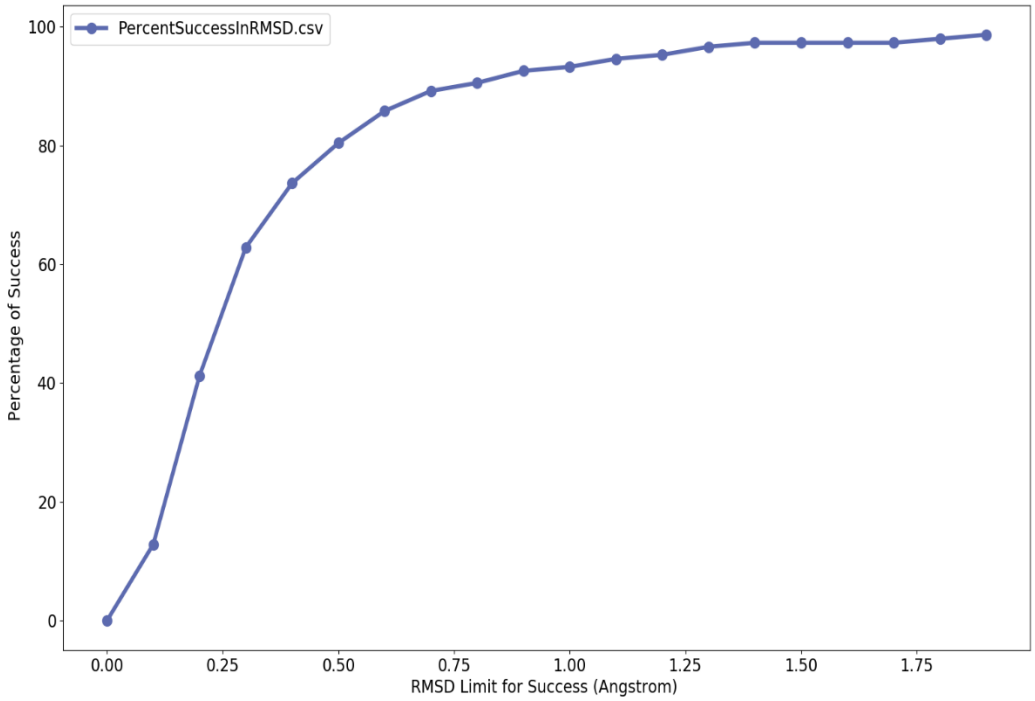
Strain energies of the top100 hits in flexible ligand docking using Vina, Vinardo, QFVina and QFVinardo scoring functions.



Mean Absolute Deviations of strain energies of the top100 docking hitlists using Vina, Vinardo, QFVina and QFVinardo scoring functions.



All of this new docking development is fundamentally based on our ability to efficiently and reliably generate high-quality QM and DFT-based conformational libraries. Our first priority is to demonstrate that these libraries are indeed of high quality. Using geometries of 150 approved drugs from the OCD database, we analyzed 25 examples from each category, with 2-7 rotatable bonds (excluding methyl groups). The figure below illustrates our success rates as a function of the heavy atom RMSD between our computed conformations and the corresponding experimental geometries.



Over 62% of drug molecules show a heavy atom RMSD below 0.3 Å, with over 80% below 0.5 Å, and approximately 93% below 1.0 Å. Achieving such high-quality conformations for thousands of molecules is made effortless with our implementation on Azure, requiring just a single command.

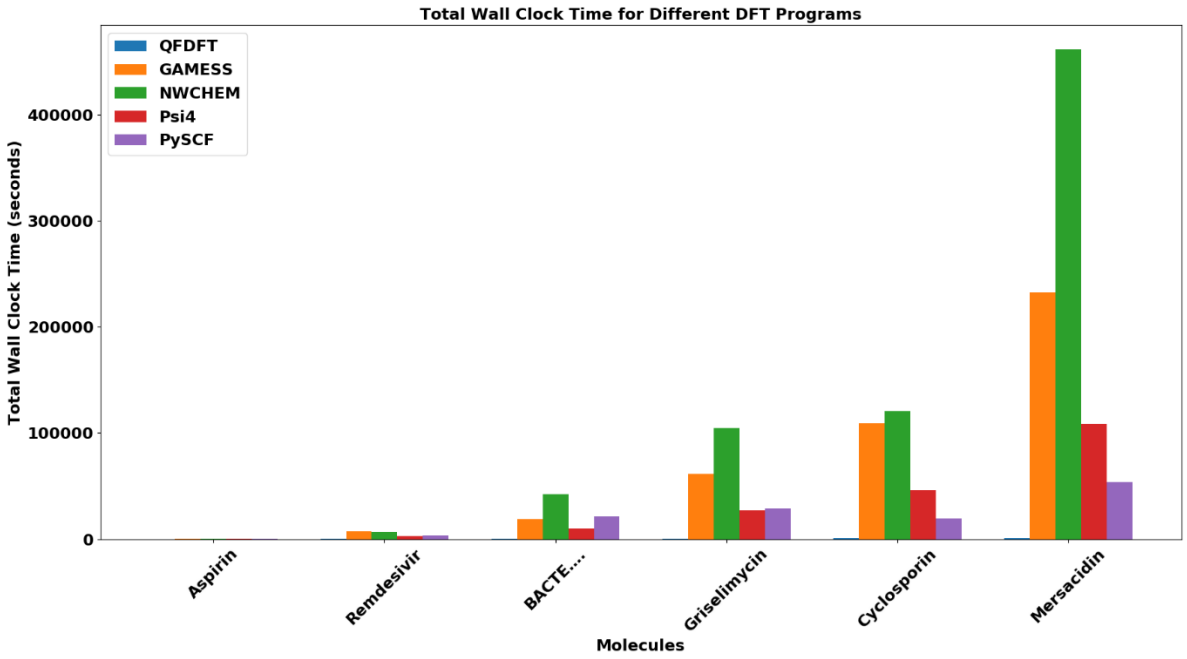
qfazurelaunch.x --azureRegion westus --azureInstanceType Standard_F16s_v2 --azureMaxSpotInstances 100 --azureProjectName MyProject qfconfsearchDFT.

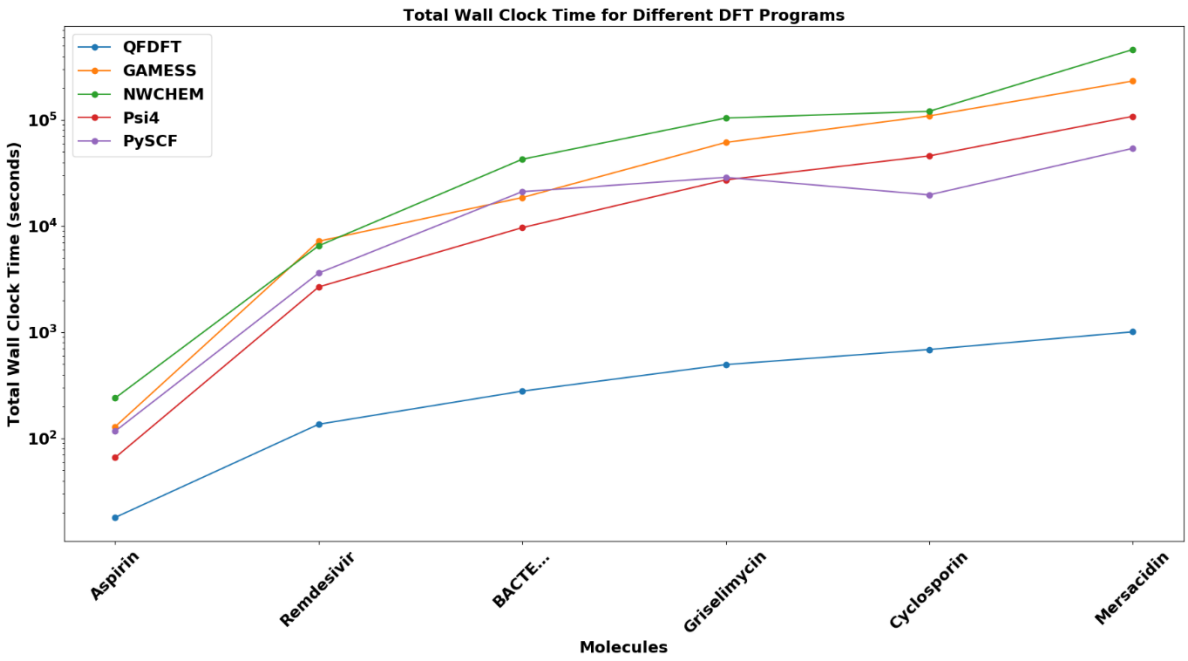
The file MyProject.tar.gz contains all the SDF files for the input molecules, and SMILES input is also supported.

This approach is made possible by our highly efficient DFT implementation, which allows us to accurately calculate ab initio conformational strain energies for use in docking scores. Using conventional DFT programs would be significantly more expensive. The next two figures illustrate the computational costs of our QFDFT compared to other widely used DFT programs.

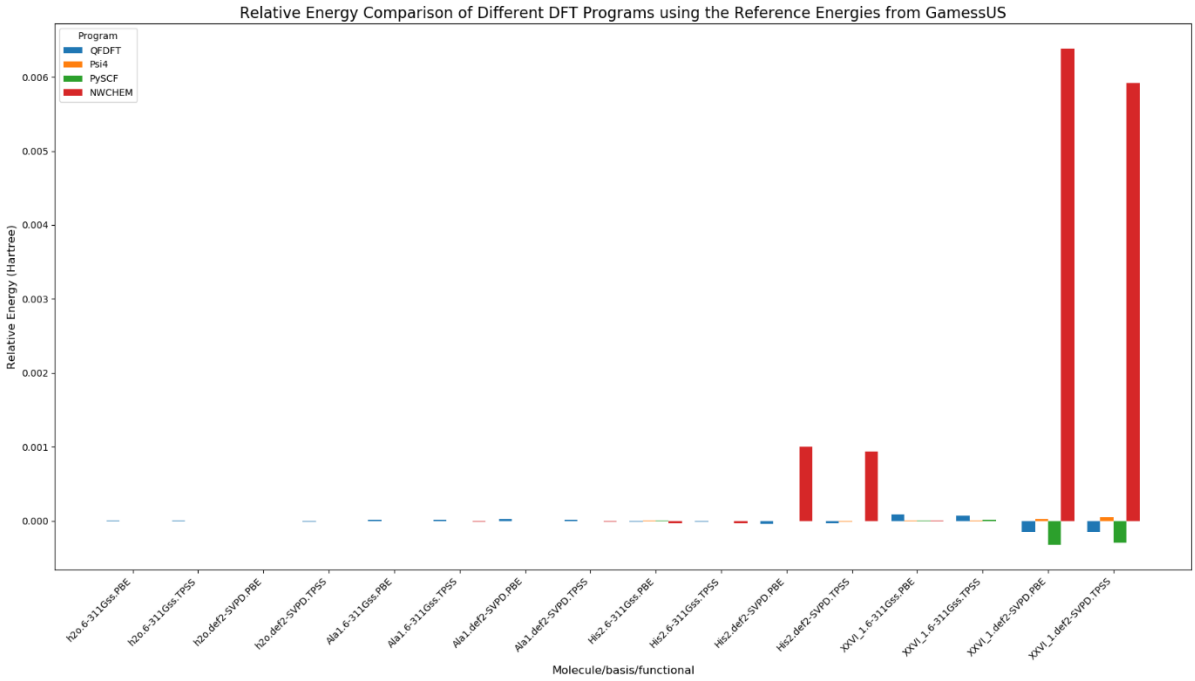
Note that the computational costs of QFDFT are barely visible in the first figure, which is why the second figure is shown on a logarithmic scale.

The DFT benchmark calculations were performed on a series of increasingly larger drug molecules using the def2-SVPD basis set and TPSS functional on an Intel i9 18-core workstation.

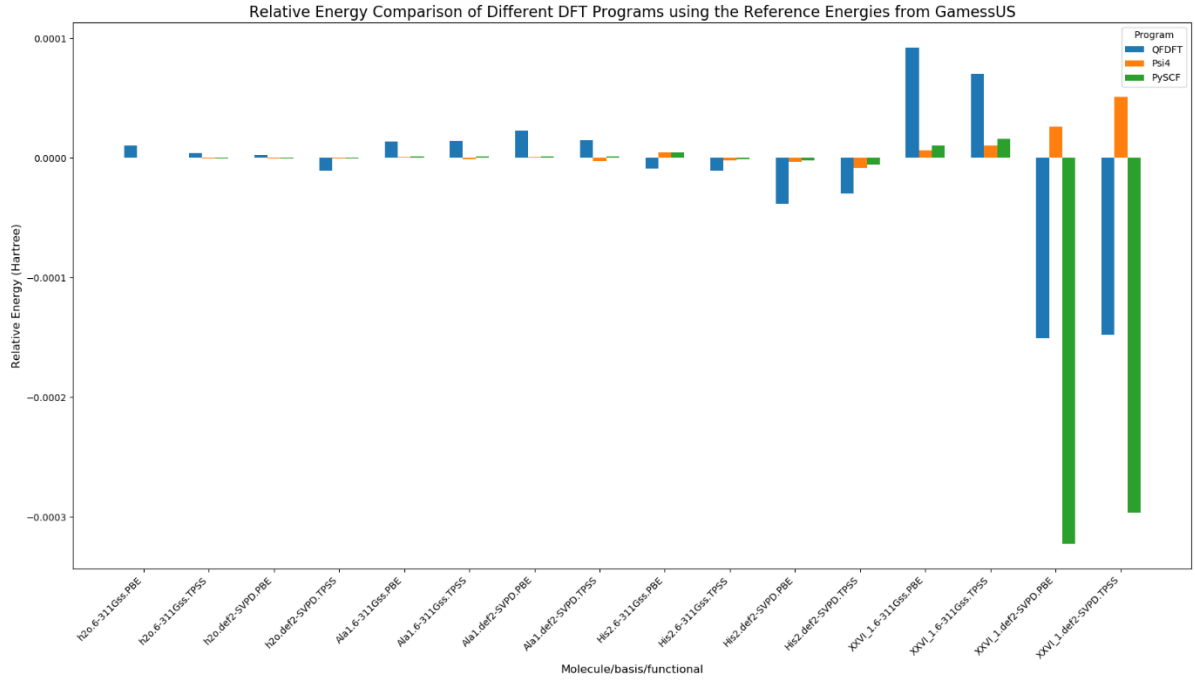




Of course, many of us have seen similar results before, but often there is a catch—such as discrepancies in the total DFT energies. This is not the case here. The next two figures compare the total DFT energies of four molecules, using all combinations of two widely used basis sets and two commonly used functionals. All energies are benchmarked against those obtained using the GAMESS-US program.



Unfortunately, NWChem really spoils the fun here with large errors, even when using oversized numerical grids. So, to keep things clean and fair, the next figure shows the same results—but without NWChem crashing the party!



Notice the small absolute values in relative energies for all three methods.

For more details on the QF docking study, see

Integrating Quantum Mechanics into Protein-Ligand Docking: Toward Higher Accuracy and Reliability | Biological and Medicinal Chemistry | ChemRxiv | Cambridge Open Engage

and stay tuned for a peer-reviewed paper. Additional information on ab initio DFT and conformational search will be published soon.