

# Comparing the expense and accuracy of methods to simulate atomic vibrations in rubrene

Makena A. Dettmann,<sup>\*,†</sup> Lucas S. R. Cavalcante,<sup>†</sup> Corina Magdaleno,<sup>†</sup> Karina Masalkovaitė,<sup>†</sup> Daniel Vong,<sup>†</sup> Jordan T. Dull,<sup>‡</sup> Barry P. Rand,<sup>‡</sup> Luke L. Daemen,<sup>¶</sup> Nir Goldman,<sup>§,†</sup> Roland Faller,<sup>†</sup> and Adam J. Moulé<sup>\*,†</sup>

<sup>†</sup>*University of California Davis, Davis, CA, USA*

<sup>‡</sup>*Department of Electrical Engineering, Princeton University, Princeton, NJ, USA*

<sup>¶</sup>*Oak Ridge National Lab, Oak Ridge, TN, USA*

<sup>§</sup>*Lawrence Livermore National Lab, Livermore, CA, USA*

E-mail: mdettmann@ucdavis.edu; amoule@ucdavis.edu

## DFT

For the DFT simulation, the VASP package was used with the opt-pbe functional, a cutoff energy of 520 eV, and a k-point sampling of 2x2x2 across all calculations. A full optimization was performed (lattice and atomic positions) with a maximum force criteria of 0.02 eV/Å. In order to calculate the phonon modes, the finite displacement method was used with a 1x1x1 supercell.

## DFTB

For the DFTB simulation, we performed the calculations with the mio1-1 Slater-Koster parameters and a similar criteria of k-point sampling and optimization as the DFT simula-

tion. Furthermore, using the same method of phonon calculation as before, we set a 2x2x2 supercell.

## DFTB/ChIMES

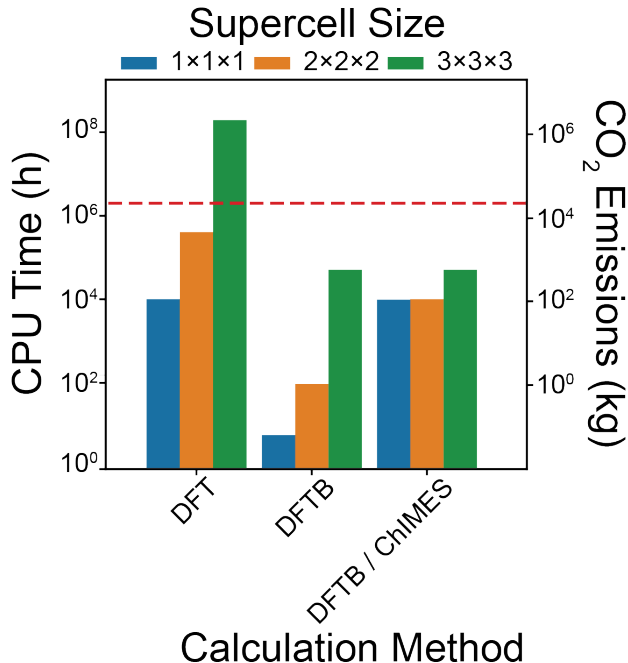


Figure 1: CPU time for the three most accurate methods for a variety of box sizes. The actual box sizes used were 1x1x1 for DFT and 2x2x2 for DFTB and DFTB/ChIMES. The other values were estimated using the scaling law for each method ( $N^3$ ). The values for DFTB with ChIMES include additional training. The dashed red line represents the yearly allocation for our group (2 million NERSC hours)

The best DFTB/ChIMES simulation began with a full structure optimization. Next a DFT-MD simulation was performed for the modeling step. VASP was with the same conditions as before, but with the PBE functional. The MD simulation was performed considering the canonical ensemble [NVT] with a trajectory integrated at 1000K with a 0.20 fs time step and the Nosé-Hoover thermostat with periodic boundary conditions. The DFT-MD simulation had in total a duration of 5 ps from which uniformly spaced frames were extracted every 100 fs producing a training set of 50 configurations. The ChIMES coefficients were fit to the training set for two body interactions up to the 8th order and

three body interactions up to the 4th order. The remaining steps of the workflow were performed using DFTB/ChIMES imposing the same set of parameters for the optimization and phonon calculation as the non-corrected DFTB simulation.

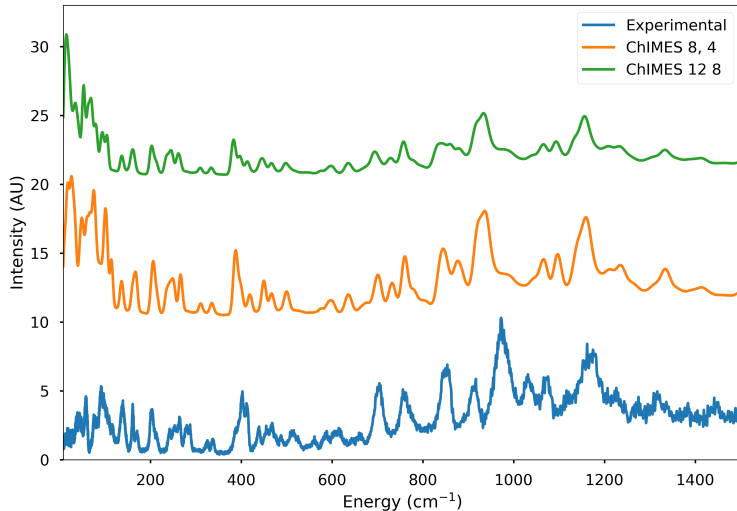


Figure 2: Comparison between two ChIMES parametrizations. The orange line was produced from a ChIMES model which used up to 8th order 2-body interactions and up to 4th order 3-body interactions. The green line was produced using a ChIMES model with up to 16th order 2-body interactions and up to 8th order 3-body interactions.

Figure 2 shows the effect of order of 2- and 3-body interactions on the predicted INS spectrum. In this case, 8th- and 4th-order 2- and 3-body interactions showed a better fit than 16th- and 8th-order interactions. This indicates that the larger order model may have been overfit. In the region between  $800\text{ cm}^{-1}$  and  $1000\text{ cm}^{-1}$ , the lower order model correctly predicts three peaks while the larger order predicts four. In this study, 2-body interaction orders were varied from 4th to 16th and 3-body interaction orders were varied from 2nd to 8th. The 8-4 model was the most successful of all.

## ML

ML was also employed to compute INS spectra. The atomistic machine learning package (AMP) along with the atomic simulation environment (ASE) was used to train a neural net to predict the forces and energies in Rubrene. The same training set used for forcebalance was used for the ML method and the Gaussian descriptor was used. First, the effects of neural net architecture were investigated. Figure 3 shows the time required to train each neural net so that the force root mean square error (RMSE) was 0.5 eV/Å. The 8x8 neural net showed the fastest convergence, so it was then trained until the maximum accuracy was achieved.

The effect of the width  $\nu$  in the Gaussian descriptor was also investigated. The width was quartered, halved, doubled, and quadrupled to determine the best width. It was found that a width twice as large as the default was able to achieve the highest convergence. The force RMSE for the best result achieved is shown in figure 4. The final RMSE was 0.34159 eV/Å. This model was used to produce the INS spectrum seen in the main paper.

## MD

In a previous publication, a method for computing an INS spectrum from an MD trajectory was described. This method was developed to be theoretically equivalent to the finite displacement supercell method paired with OCLIMAX. This study extends that discussion by examining the effect of forcefield fitting on predicted INS spectra.

The initial forcefield was taken from the automated topology builder (ATb). It was based on a modified GROMOS 54A7 forcefield for gromacs. The .pdb file was also taken from ATb. In order to transform the molecule in a vacuum into a periodic crystal, a periodic .cif file from the Cambridge Crystallographic Data Centre was employed. ASE was used to place the single molecule inside the crystal lattice. The single molecule was then translated and rotated until it matched the lattice molecule as closely as possible. This was repeated for

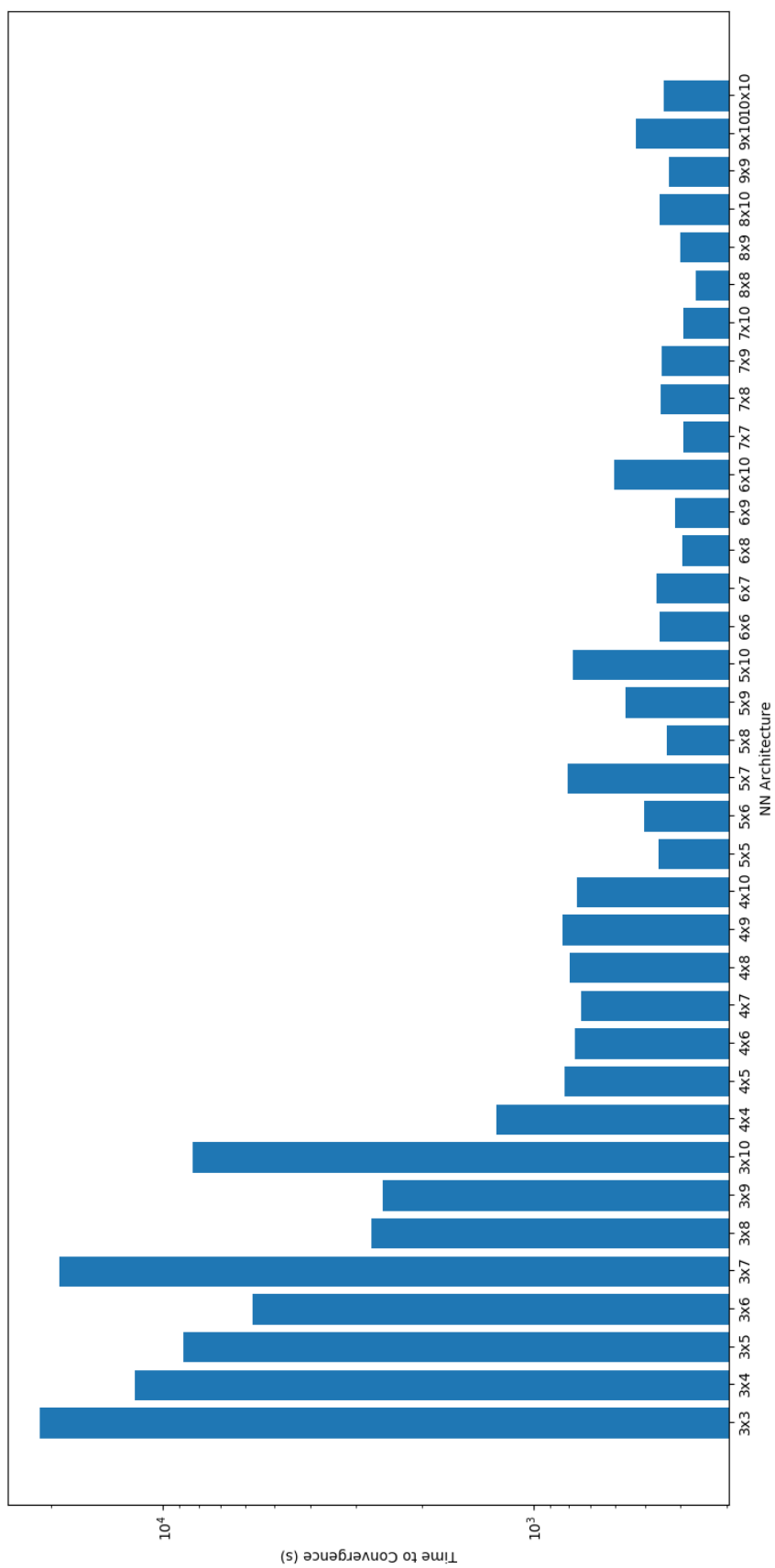


Figure 3: Convergence time for different 2-layer neural net architectures. All models were fit to within 0.5 eV/Å for the force RMSE. Here, a  $3 \times 10$  architecture is a neural net with 3 nodes in the first hidden layer and 10 nodes in the second hidden layer.

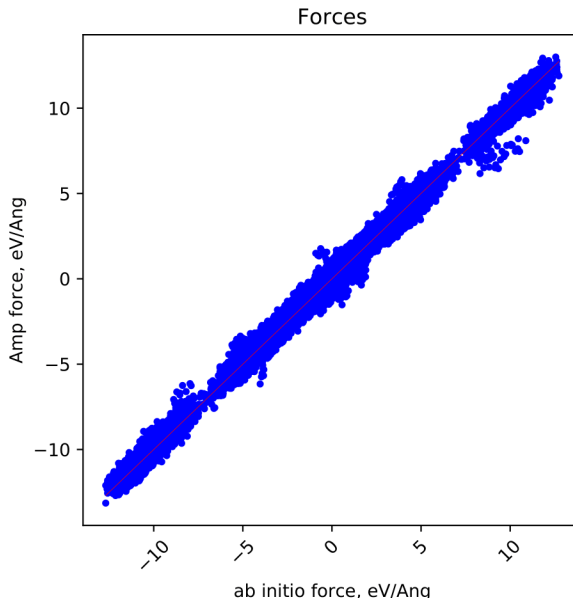


Figure 4: The force RMSE of the ML model compared to VASP. An 8x8 architecture with a Gaussian descriptor in which  $\eta$  values were double the defaults.

each molecule in the unit cell. Once these positions were determined, the molecules needed to be untwisted. The molecule as downloaded from ATb was twisted around its backbone (Figure 5a), but crystalline rubrene (Figure 5b) is not. Next, a 4x2x1 supercell was made and simulated annealing was performed. The final structure is shown in Figure 5c.

After creating the supercell, the system was equilibrated to 10 Kelvin. Then an NVT training MD simulation was performed in gromacs. The training run was 500ps and the positions were extracted every 0.5ps which generated 1000 configurations. VASP was then used to calculate the forces and energies at each configuration. Then forcebalance was used to improve the forcefield. A number of different training sessions were performed in which different parameters were allowed to change. Allowing all bonds, angles, dihedrals, and Lennard-Jones parameters to change yielded a similar result as just allowing the Lennard-Jones parameters to change (Figure 6). Thus, a forcefield with only improved Lennard-Jones parameters was used.

After the best forcefield was produced, the box effect of the box size was also investigated. Figure 7 shows how the supercell size affected the final INS spectrum. Supercells up to

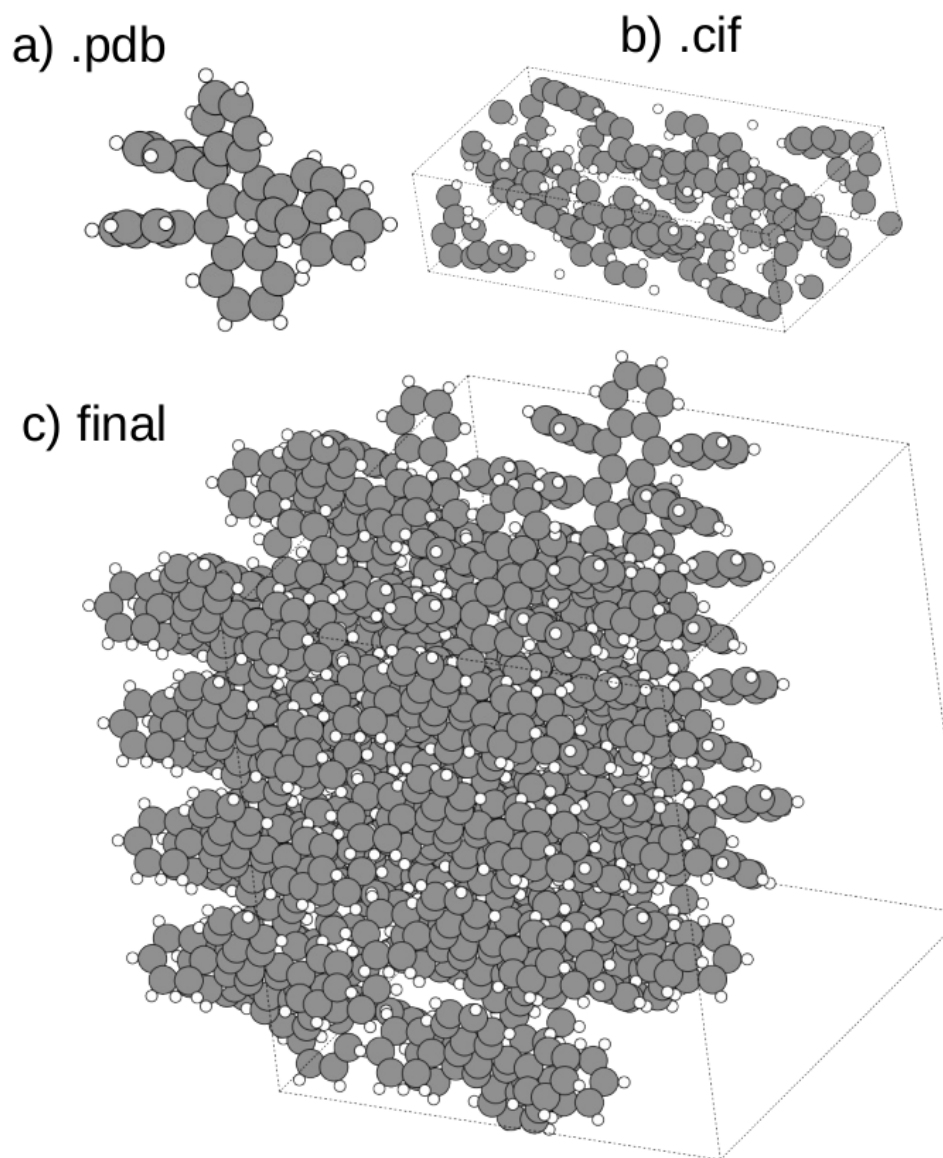


Figure 5: Atomic structures of rubrene. a shows the structure from the .pdb file. B shows the structure from a .cif file and c shows the final structure after transforming a to look like b, creating a supercell, and performing a simulated annealing run.

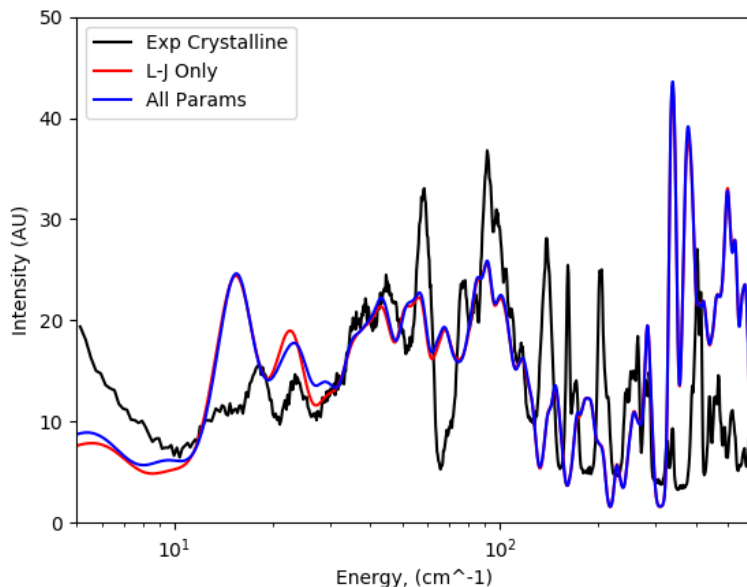


Figure 6: Effect of trained parameters on INS spectrum. L-J Only indicates that only the Lennard-Jones parameters were trained and All Params means that every coefficient in the forcefield was allowed to change (various  $k$  values along with  $\sigma$  and  $\epsilon$  in the main paper). A log scale is used because differences are subtle on a linear scale.

4x16x8 were used to compute an INS spectrum. Larger supercells were not investigated due to computational limitations. The results show that a larger simulation size produces more accurate results and supercell size is important in these types of simulations.

Finally, the velocity autocorrelation method was compared to the finite displacement supercell method. Figure 8 shows the INS spectra computed using each method. The two spectra are largely similar with the primary difference in the low energy region. This is likely because a small box size was required for the finite displacement method while a larger box was used for the velocity autocorrelation method.

## Times

The CPU times required to produce the spectra shown in Figure 3 of the main article are shown in Figure 9. For MD and DFTB/ChIMES times including training and without



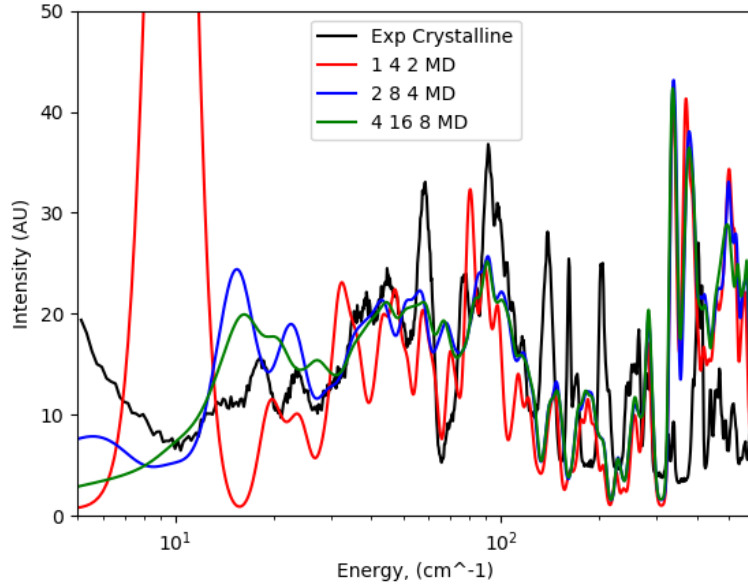


Figure 7: Effect of the production run's box size on the final INS spectrum. 1 4 2 MD means that a  $1 \times 4 \times 2$  supercell was used, 2 8 4 MD indicates that a  $2 \times 8 \times 4$  supercell was used, and 4 16 8 MD indicates that a  $4 \times 16 \times 8$  supercell was used. A log scale is used for the x-axis because the majority of the differences are at low energies.

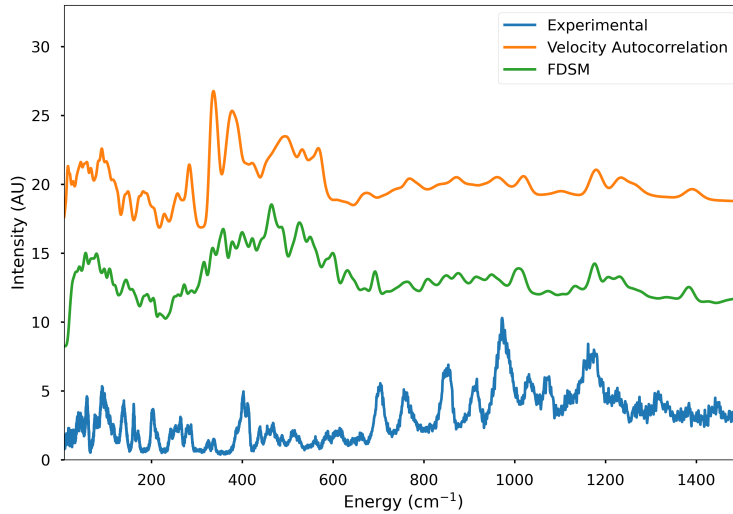


Figure 8: The INS spectrum computed using the finite displacement supercell method (green) compared to the INS spectrum computed using the velocity autocorrelation method (orange) and the experimental INS spectrum (blue).

training are shown. This is because these methods should be transferrable to other systems. The time without training is not shown for ML because it is not known how well a neural net potential trained on one system would perform with a different system. The color indicates the supercell size used to compute the spectra in Figure 3 of the main article.

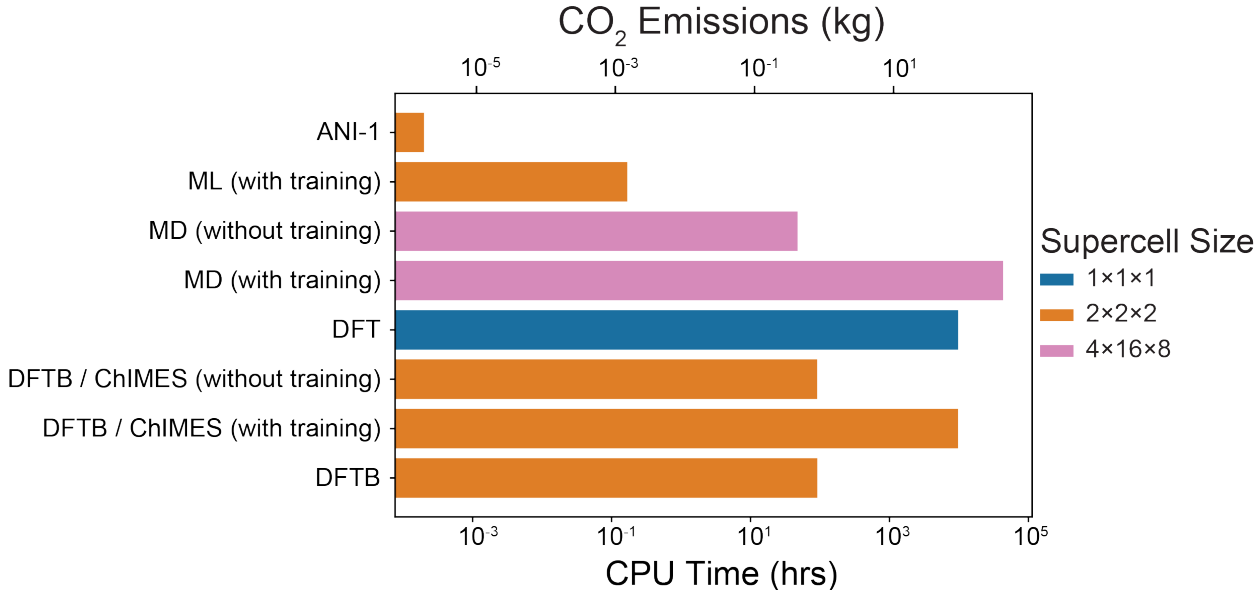


Figure 9: The times and estimated CO<sub>2</sub> emissions for each method. These times correspond to the best spectrum produced by each method. These spectra are the ones plotted in figure 3 of the main article. The supercell sizes used for each calculation are indicated by color. For MD and DFTB/ChIMES, times are shown both with and without training.