SUPPORTING INFORMATION

Evaluating Computational Shortcuts in Supercell-Based Phonon Calculations of Molecular Crystals: The Instructive Case of Naphthalene

Tomas Kamencek[1,2], Sandro Wieser[1], Hirotaka Kojima[3], Natalia Bedoya-Martínez[4], Johannes P. Dürholt[5], Rochus Schmid[5], and Egbert Zojer[1]

[1]Institute of Solid State Physics, Graz University of Technology, NAWI Graz, Petersgasse 16, 8010 Graz, Austria

[2]Institute of Physical and Theoretical Chemistry, Graz University of Technology, NAWI Graz, Stremayrgasse 9, 8010 Graz, Austria

[3]Division of Materials Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan

[4]Materials Center Leoben, Roseggerstraße 12, 8700 Leoben, Austria

[5]Chair of Inorganic Chemistry 2, CMC Group, Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum, Germany

1. Data availability

Input files for the used codes, geometries and harmonic force constants are available online in the following data repository:

https://github.com/cmc-rub/supporting_data/tree/master/85-Kamencek_JCTC_2020

2. Importance of sampling phonons in the entire reciprocal space for calculating thermodynamic properties

In spite of the practice common in literature to neglect phonon dispersions when calculating thermodynamic properties of molecular crystals, non-Γ-phonons can have a significant impact on the results. Fig. S1 shows how the evolutions of Helmholtz free energy and the heat capacity differ, when considering only Γ-phonons or phonons from the entire first Brillouin zone (sampled on a 9×10×9 mesh; see main text). For the free energy one observes an energy difference of more than 0.1 eV per unit cell at room temperature. Contrary, the heat capacity is mainly influenced at low temperatures: if only Γ-phonons are considered, there is a non-vanishing contribution at zero frequency giving rise to a violation of the third law of thermodynamics.

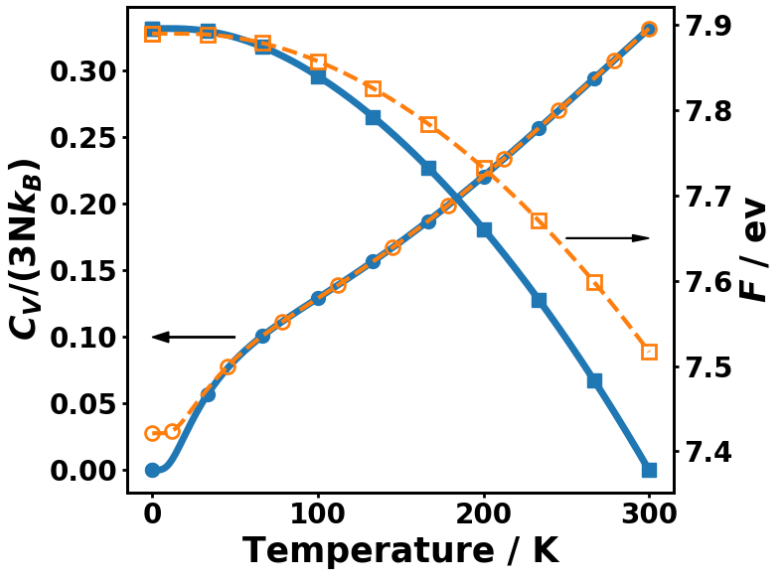

Fig. S1: *Phonon heat capacity C_V normalized by the number of vibrational modes 3N (circles) and vibrational free energy F (squares) as a function of temperature calculated from phonons of the entire first Brillouin zone (blue solid lines) and from Γ-phonons only (orange dashed lines).*

3. Raman spectra simulation

The intensities I_k associated with the k^{th} vibrational mode is related to the Raman activity A_k according to the following equations [1]. Besides the Raman activity, the calculated (Stokes) intensity depends on the fourth power of the frequency difference between the vibrational mode and the excitation radiation as well as a thermal occupation factor including the Bose-Einstein distribution *n* (eq. (1a)).

The Raman activity A_k depends on the geometry of the experimental setup and on the sample. This is considered in eq. (1b) for the assumption of the incident and the reflected light beam being orthogonal to each other. For isotropic and homogeneous samples, the Raman activity can be written according to eq. (1b) with the occurring quantities being defined in eq. (1c) and (1d).

$$I_k \propto \frac{(\Omega_{in} - \omega)^4}{\omega_k} \; (n(\omega, T) + 1) \; A_k \tag{1a}$$

$$A_k = 45\,\gamma_k^2 + 7\beta_k^2 \tag{1b}$$

$$\gamma_k = \frac{1}{3}\sum_{i=1}^{3} \chi_{ii,k} \tag{1c}$$

$$\beta_k^2 = \frac{1}{2}\Big\{\big(\chi_{11k} - \chi_{22,k}\big)^2 + \big(\chi_{11k} - \chi_{33,k}\big)^2 + \big(\chi_{33k} - \chi_{22,k}\big)^2 + 6\big(\chi_{12,k}^2 + \chi_{23,k}^2 + \chi_{13,k}^2\big)\Big\} \quad (1d)$$

$$\chi_{ij,k} = \frac{\partial \alpha_{ij}}{\partial Q_k} = \sum_{l=1}^{3N} \frac{\partial \alpha_{ij}}{\partial u_l} \frac{u_k^{(l)}}{\left\| u_k^{(l)} \right\|} \quad (1e)$$

$$u_k^{(l)} = \frac{e_k^{(l)}}{\sqrt{m_l}}; \quad \left\| u_k^{(l)} \right\|^2 = \sum_{l=1}^{3N} \left(u_k^{(l)} \right)^2 \quad (1f)$$

The most relevant quantities appearing in these expressions are the derivatives of the polarization tensor α_{ij} with respect to normal mode coordinates Q_k. They constitute the Raman tensor, $\chi_{ij,k}$. The Raman tensor can either be calculated directly by displacing the geometry along the normal mode coordinates and calculating the change in polarizability (or electric susceptibility) as a function of the normal mode displacement.

Alternatively, one can rewrite the derivative with respect to the normal mode coordinate Q_k as a derivative with respect to the cartesian displacement u_l. The associated transformation is then given by the normalized eigendisplacements $u_k^{(l)}$, *which can be understood as the* mass weighted phonon eigenvectors (polarization vectors) $e_k^{(l)}$. The normalization of the eigendisplacement vectors is the common vector norm extended to 3*N* dimensions. Although the direct approach is more useful when calculating Raman tensors for specific modes, the approach based on cartesian displacements is much more efficient for systems with large number of symmetries since many cartesian derivatives can be obtained from simple symmetry transformations, and, thus, the number of symmetry-inequivalent displacements necessary to simulate Raman tensors for all modes can be drastically reduced.

Practically, the dielectric function was calculated with *VASP* applying density functional perturbation theory for each displaced geometry produced by *PHONOPY*. The symmetry-irreducible Raman tensors were calculated from eq. (1e), while the remaining ones were obtained by applying the respective point group symmetry operations to those rank 3 tensors.

The Raman activities of the isolated naphthalene molecule were calculated with the *Gaussian 16* package (Revision A.03) [2] after a proper geometry optimization employing the D3-BJ van-der-Waals *a posteriori* correction. For both used functionals (PBE and B3LYP) we employed the triple-zeta Gaussian-type basis set 6-311++G(d,p) including diffuse and polarization functions. Subsequently, the Raman activities (using the calculation type identifiers `Opt` and `Freq`) were calculated using the equations above with the fully automatic routines of *Gaussian*.

The plotted spectra consist of Lorentzian function with a full-width-at-half-maximum of 0.2 THz placed at each resonance.

4. Converging DFT settings

3.1 Used Pseudopotentials and global setting

The following *VASP* pseudopotentials were used for hydrogen and carbon, respectively: PAW_PBE H 15Jun2001, PAW_PBE C 08Apr2002

Additionally, the following simulation settings were used throughout all tests:

```
LREAL =.FALSE.;
ALGO = Fast;
ISMEAR = 0;
SIGMA = 0.05;
```

3.1 Impact of DFT settings on phonon frequencies

Especially the low frequency phonon bands are often found to be relatively sensitive to the simulation parameters, so that tight convergence criteria must be chosen, consuming a high amount of computational resources. We, thus, studied how the *VASP*-specific parameters controlling the plane wave energy cutoff (`ENCUT`), the SCF convergence criterion (`EDIFF`) and a global precision setting (`PREC`) influence the resulting Γ frequencies. The geometry, which was optimized with the thoroughly converged parameters described in the main text, was kept the same for all tests. The root-mean-square (RMS) error and the maximum absolute deviation of the Γ frequencies with respect to the reference simulation were recorded. The associated maps can be seen Fig. S2. We find that the errors depend much more strongly on the energy cutoff than on the SCF convergence criterion. Furthermore, the error does not monotonically decrease with the cutoff. Interestingly, for smaller cutoffs, choosing normal precision results in lower RMS errors than the accurate setting. The supposedly most underconverged settings (top left in Fig. S2) also result in smaller RMS errors than calculations with the same energy cutoff but different SCF convergence criteria

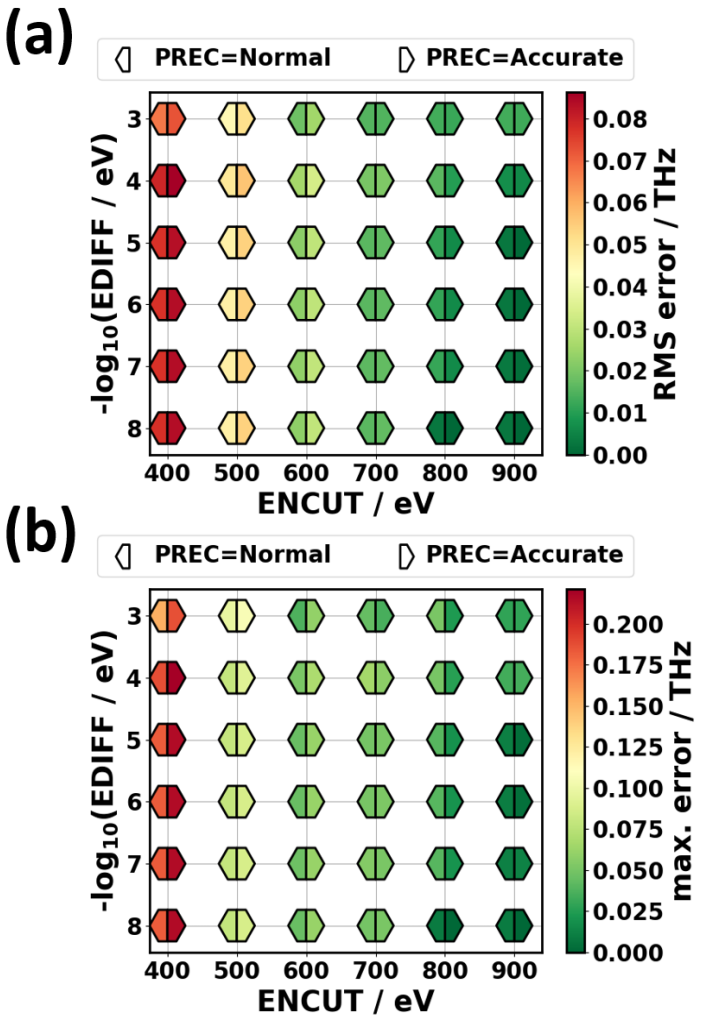

Fig. S2: *Accuracy map of Γ–frequencies dependent on the three VASP parameters* `PREC`, `ENCUT`, `EDIFF`, *controlling the global precision settings, the plane wave energy cutoff, and the SCF convergence criterion, respectively. The (a) root-mean-square (RMS) error and the (b) maximum error were calculated with respect to the reference calculation (*`PREC` *=Accurate,* `ENCUT` *=900 eV,* `ENCUT` *=10^{-8} eV).*

Although the gain in computing time for the most underconverged settings is not extra-ordinarily large (about a factor 3.7 given our computing architecture Intel Xeon E5-25650 CPUs, the used level of parallelization and our compilation of the code), it is still instructive to compare the phonon bands obtained with the most "economic" settings to the most accurate reference results. We base the comparison of the band structure on the low frequency modes (below 9 THz ≈ 300 cm^{-1}) due to the reasons given in the Methodology section of the main manuscript. The band structure and the density of states (DOS) obtained with the "economic" settings are compared to the reference calculation in Fig. S3. Although the results are supposedly highly underconverged, the agreement in the bands and

especially in the DOS is surprisingly good. Only the width of the band gap between ~4.1-5.3 THz slightly increases, as the bands at the lower edge of the gap are somewhat shifted to lower frequencies At ~2 THz and ~3 THz at the A-point some further small differences can be seen. Notably, the lowest acoustic band (transverse acoustic) in ΓA direction is the only acoustic band, whose band width is notably underestimated. The most pronounced difference is the band dispersion of the second lowest band along XA which is much flatter in "DFT eco" than in the reference calculation "DFT ref".

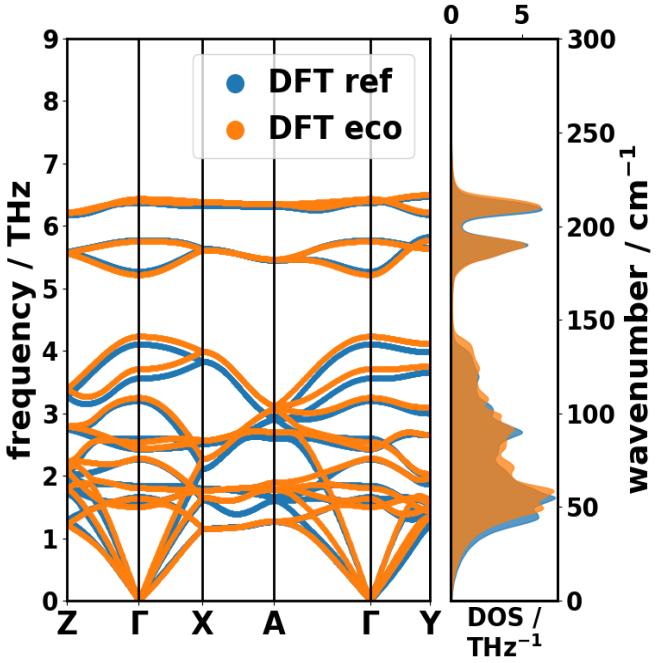

Fig. S3: *Phonon band structure and DOS of (non-deuterated) naphthalene obtained with DFT+D3-BJ. Blue: accurate reference calculation. Orange: underconverged ("economic") DFT settings*

To quantify the deviation, we sampled the 1BZ at 125 **q**-points and calculated the RMS deviation between the "economic" and the reference data. When considering all phonons of the material, an RMS deviation of 0.08 THz is obtained, which ich only slightly larger than the calculated RMS deviation, when considering only Γ-point frequencies (0.07 THz). The RMS deviation for bands up to frequencies of 9 THz is slightly increased (0.11 THz), which suggests that the reduced simulation accuracy more severely impacts the low frequency modes of mostly inter-molecular character. The higher-lying, intra-molecular modes are apparently less affected.

5. D3-BJ parameters used in DFTB and DFT

The D3-BJ van der Waals (vdW) correction depends on four global parameters (a_1, a_2, s_6, and s_8) with s_6 usually being kept fixed at unity. The other three parameters were chosen according to the

suggested standard values provided in the respective user manuals [3],[4]. In this work, we used the recommended standard parameters for the PBE and DFTB3 functional, respectively.

Tab. S1: *Standard D3-BJ parameters used for the given functionals according to the VASP and DFTB+ manuals.*

Functional	a_1 / Bohr	a_2	s_6	s_8
DFTB3	0.746	4.191	1.0	3.209
PBE	0.4289	4.4407	1.0	0.7875

6. Cell optimization and unit cell rescaling in DFTB

Unlike *VASP*, which allows to optimize unit cell parameters within the constraint of constant volume to perform a fit to an equation of state, the used version of *DFTB+* (version 18.1) does not provide this functionality. To overcome this problem, a different approach was chosen for optimizing the unit cell while keeping the type of Bravais lattice (simple monoclinic lattice with four lattice parameters: the three lengths of lattice vectors *a*, *b*, and *c* as well as the monoclinic angle β): for a set of fixed monoclinic angles, the lattice vector lengths (together with the atomic coordinates) were optimized with fixed angles. Afterwards the optimum monoclinic angle was obtained by fitting a parabola to the energy-vs.-angle data (see Fig. S4). This optimum angle was used in a last step for a unit cell, whose remaining lattice vectors and atomic positions were optimized to end up with the final fully optimized geometry.

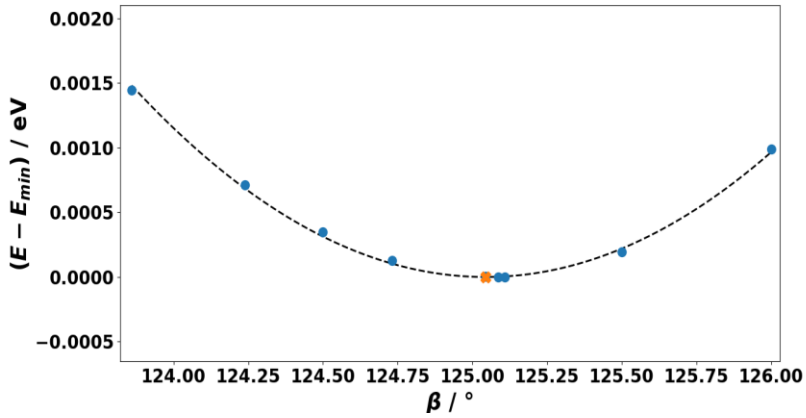

Fig. S4: *Total energy difference with respect to the minimum as a function of the monoclinic angle β calculated with DFTB+. The blue circles correspond to optimized unit cells with fixed monoclinic angle, the orange cross marks the fitted minimum.*

7. Unit cell rescaling in DFTB

In order to find the optimal rescaled unit cell volume in the DFTB calculations, we calculated the root mean square deviations (RMSD) in frequencies compared to the DFT/D3-BJ reference (see main text). Regardless of whether only the Γ-point phonons or modes from the entire first Brillouin zone are used, the scaling factor which minimizes the RMSD of frequencies was found to be 0.95 %. Note that for each volume, the atomic positions were fully optimized.

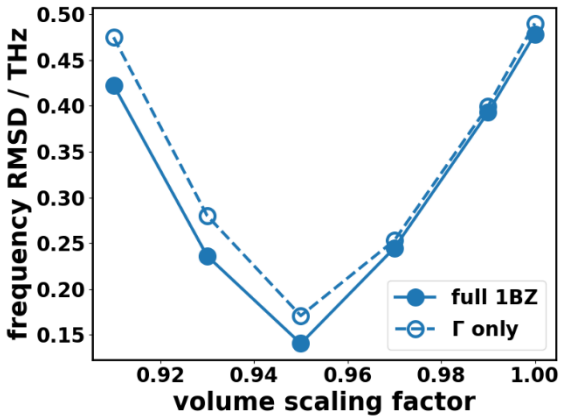

Figure S5. *Root-mean-square deviation (RMSD) of frequencies between DFTB calculations and the DFT/D3-BJ reference. The DFTB frequencies have been obtained for a DFT/D3-BJ unit cell, whose volume has been rescaled by the factor given on the horizontal axis (for details see main text). The solid line with filled symbols represents the deviations calculated for the entire first Brillouin zone, while for the dashed line with empty symbols Γ–point frequencies have been considered. Note that the DFTB energy of the "optimally" rescaled unit cell is ~30 meV higher than for the DFTB-optimized cell and shows a tensile stress of ~ -5 kbar.*

8. MOF-FF parametrization

The MOF-FF functional form is built as a sum over many independent contributions [5]:

$$E = E_{str} + E_{bnd} + E_{tor} + E_{oop} + E_{coul} + E_{vdW} \qquad (2a)$$

Here E_{str} is the stretch bond potential, E_{bnd} is the bending potential, E_{tor} is the torsional potential, E_{oop} is the out of plane bending potential, E_{coul} is the electrostatic potential and E_{vdW} is the van der Waals potential.

Bonded interaction terms:

$$E_s^{str} = \frac{1}{2}k_s\left(r_s - r_s^{ref}\right)^2\left[1 - 2.55\left(r_s - r_s^{ref}\right) + \frac{7}{12}\left(2.55\left(r_s - r_s^{ref}\right)\right)^2\right] \quad (2b)$$

$$E_b^{bnd} = \frac{1}{2}k_b\left(\theta - \theta_b^{ref}\right)^2\left[1 - 0.014\left(\theta_b - \theta_b^{ref}\right) + 5.6\cdot 10^{-5}\left(\theta_b - \theta_b^{ref}\right)^2 - 7\right. \quad (2c)$$
$$\left.\cdot 10^{-7}\left(\theta_b - \theta_b^{ref}\right)^3 + 2.2\cdot 10^{-8}\left(\theta_b - \theta_b^{ref}\right)^4\right]$$

$$E_t^{tor} = \sum_n \frac{V_t^n}{2}[1 + \cos(n\phi_t + \phi_t^n)] \quad (2d)$$

$$E_o^{oop} = \frac{1}{2}k_o(\theta_o - \theta_0)^2 \quad (2e)$$

Cross-terms between bonded interactions between atoms of a bending angle:

$$E_b^{str-bnd} = \left(\theta_b - \theta_b^{ref}\right)\left[k_{sb1}\left(r_{b1} - r_{b1}^{ref}\right) + k_{sb2}\left(r_{b2} - r_{b2}^{ref}\right)\right] \quad (2f)$$

$$E_b^{str-str} = k_{ss}\left(r_{b1} - r_{b1}^{ref}\right)\left(r_{b2} - r_{b2}^{ref}\right) \quad (2g)$$

Torsional cross-term, that was not originally contained in MOF-FF (underlined in the following input file):

$$E_{bb13} = N\left(r_{ij} - r_1\right)\left(r_{kl} - r_3\right) \quad (2h)$$

Non-bonded interaction terms:

$$E_{ij}^{vdW} = \varepsilon_{ij}\left\{1.85\cdot 10^5 \exp\left(-12\frac{d_{ij}}{d_{ij}^0}\right) - 2.25\left(\frac{d_{ij}^0}{d_{ij}}\right)^6\left[1 + 6\left(\frac{0.25 d_{ij}^0}{d_{ij}}\right)^{14}\right]^{-1}\right\} \quad (2i)$$

$$E_{ij}^{coul} = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{d_{ij}}\mathrm{erf}\left(\frac{d_{ij}}{\sigma_{ij}}\right) \quad (2j)$$

k, α, V ... fitted parameters

r ... interatomic distance

θ ... bending angle

ϕ ... torsional angle

d_{ij} ... non-bonded atomic distance

ϵ_{ij} ... van der Waals potential well depth

q_i ... atomic point charge

σ_{ij} ... Gaussian charge distribution width

The reference data obtained with the Turbomole [6] software package (version 7.3) was obtained utilizing the PBE functional [7], employing the D3-BJ [8],[9] dispersion correction and using the def2-TZVPP [10] basis set. The SCF convergence criterion was set to 10^{-7} Hartree (~$2 \cdot 10^{-6}$ eV). The atomic position was optimized until a maximum force of 10^{-3} Hartree/Bohr (~$5 \cdot 10^{-2}$ eV/ Å) and an energy convergence of 10^{-6} Hartree (~$2 \cdot 10^{-5}$ eV) was reached.

The parameters for the force field potential used for obtaining vibrational properties can be found in form of a LAMMPS input file. The syntax of the parameters can be found in the official manual: https://lammps.sandia.gov/doc/Manual.html

```
# MOF-FF Potential parameters

# Explanation of atom descriptors:

# c3_c2h1: atom of species c bonded to three different atoms,

#          two of which are of species c and one of species h

# c3_c2h1S: the S indicates the atoms closer to the molecule center

pair_style buck6d/coul/gauss/long 0.9000     0.9000    12.0000


pair_coeff     1     1           10304       3.0612245       457.17971       4.5218516
0.60800971    # buck6d->(c3_c2h1@naph)|naphthalene/gaussian->(c3_c2h1@naph)|naphthalene <-->
buck6d->(c3_c2h1@naph)|naphthalene/gaussian->(c3_c2h1@naph)|naphthalene

pair_coeff     1     2           10304       3.0612245       457.17971       4.5218516
0.60800971    # buck6d->(c3_c2h1@naph)|naphthalene/gaussian->(c3_c2h1@naph)|naphthalene <-->
buck6d->(c3_c2h1S@naph)|naphthalene/gaussian->(c3_c2h1S@naph)|naphthalene

pair_coeff     1     3           10304       3.0612245       457.17971       4.5218516
0.60800971    # buck6d->(c3_c2h1@naph)|naphthalene/gaussian->(c3_c2h1@naph)|naphthalene <-->
buck6d->(c3_c3@naph)|naphthalene/gaussian->(c3_c3@naph)|naphthalene

pair_coeff     1     4       6157.8178       3.4682081       129.19572      0.78772886
0.73006542    # buck6d->(c3_c2h1@naph)|naphthalene/gaussian->(c3_c2h1@naph)|naphthalene <-->
buck6d->(h1_c1@naph)|naphthalene/gaussian->(h1_c1@naph)|naphthalene

pair_coeff     1     5       6157.8178       3.4682081       129.19572      0.78772886
0.73006542    # buck6d->(c3_c2h1@naph)|naphthalene/gaussian->(c3_c2h1@naph)|naphthalene <-->
buck6d->(h1_c1S@naph)|naphthalene/gaussian->(h1_c1S@naph)|naphthalene

pair_coeff     2     2           10304       3.0612245       457.17971       4.5218516
0.60800971    # buck6d->(c3_c2h1S@naph)|naphthalene/gaussian->(c3_c2h1S@naph)|naphthalene <-->
buck6d->(c3_c2h1S@naph)|naphthalene/gaussian->(c3_c2h1S@naph)|naphthalene
```

```
pair_coeff     2     3           10304       3.0612245       457.17971       4.5218516
0.60800971    # buck6d->(c3_c2h1S@naph)|naphthalene/gaussian->(c3_c2h1S@naph)|naphthalene <-->
buck6d->(c3_c3@naph)|naphthalene/gaussian->(c3_c3@naph)|naphthalene

pair_coeff     2     4       6157.8178       3.4682081       129.19572      0.78772886
0.73006542    # buck6d->(c3_c2h1S@naph)|naphthalene/gaussian->(c3_c2h1S@naph)|naphthalene <-->
buck6d->(h1_c1@naph)|naphthalene/gaussian->(h1_c1@naph)|naphthalene

pair_coeff     2     5       6157.8178       3.4682081       129.19572      0.78772886
0.73006542    # buck6d->(c3_c2h1S@naph)|naphthalene/gaussian->(c3_c2h1S@naph)|naphthalene <-->
buck6d->(h1_c1S@naph)|naphthalene/gaussian->(h1_c1S@naph)|naphthalene

pair_coeff     3     3           10304       3.0612245       457.17971       4.5218516
0.60800971    # buck6d->(c3_c3@naph)|naphthalene/gaussian->(c3_c3@naph)|naphthalene <-->
buck6d->(c3_c3@naph)|naphthalene/gaussian->(c3_c3@naph)|naphthalene

pair_coeff     3     4       6157.8178       3.4682081       129.19572      0.78772886
0.73006542    # buck6d->(c3_c3@naph)|naphthalene/gaussian->(c3_c3@naph)|naphthalene <-->
buck6d->(h1_c1@naph)|naphthalene/gaussian->(h1_c1@naph)|naphthalene

pair_coeff     3     5       6157.8178       3.4682081       129.19572      0.78772886
0.73006542    # buck6d->(c3_c3@naph)|naphthalene/gaussian->(c3_c3@naph)|naphthalene <-->
buck6d->(h1_c1S@naph)|naphthalene/gaussian->(h1_c1S@naph)|naphthalene

pair_coeff     4     4            3680               4          32.805      0.10690769
0.9771554    # buck6d->(h1_c1@naph)|naphthalene/gaussian->(h1_c1@naph)|naphthalene <-->
buck6d->(h1_c1@naph)|naphthalene/gaussian->(h1_c1@naph)|naphthalene

pair_coeff     4     5            3680               4          32.805      0.10690769
0.9771554    # buck6d->(h1_c1@naph)|naphthalene/gaussian->(h1_c1@naph)|naphthalene <-->
buck6d->(h1_c1S@naph)|naphthalene/gaussian->(h1_c1S@naph)|naphthalene

pair_coeff     5     5            3680               4          32.805      0.10690769
0.9771554    # buck6d->(h1_c1S@naph)|naphthalene/gaussian->(h1_c1S@naph)|naphthalene <-->
buck6d->(h1_c1S@naph)|naphthalene/gaussian->(h1_c1S@naph)|naphthalene


bond_style hybrid class2 morse harmonic


bond_coeff     3 class2     1.096263   366.570728  -934.755355  1390.448591    # mm3-
>(c3_c2h1@naph,h1_c1@naph)|naphthalene

bond_coeff     6 class2     1.433161   381.108715  -971.827224  1445.592996    # mm3-
>(c3_c3@naph,c3_c3@naph)|naphthalene

bond_coeff     1 class2     1.391210   490.285107 -1250.227022  1859.712695    # mm3-
>(c3_c2h1@naph,c3_c2h1S@naph)|naphthalene

bond_coeff     2 class2     1.454449   349.932367  -892.327535  1327.337208    # mm3-
>(c3_c2h1@naph,c3_c2h1@naph)|naphthalene
```

```
bond_coeff     4 class2     1.435896   383.965503  -979.112033  1456.429149    # mm3-
>(c3_c2h1S@naph,c3_c3@naph)|naphthalene

bond_coeff     5 class2     1.090635   382.608333  -975.651250  1451.281234    # mm3-
>(c3_c2h1S@naph,h1_c1S@naph)|naphthalene


angle_style hybrid class2/p6 cosine/buck6d


angle_coeff     8 class2/p6      108.146325    94.230110   -75.585826    17.322995   -
12.406682    22.341015    # mm3->(c3_c2h1S@naph,c3_c3@naph,c3_c3@naph)|naphthalene

angle_coeff     8 class2/p6 bb    61.790580     1.435896     1.433161

angle_coeff     8 class2/p6 ba    69.044622    24.678564     1.435896     1.433161

angle_coeff     7 class2/p6      120.083549    50.819349   -40.764279     9.342485    -
6.691062    12.048759    # mm3->(c3_c2h1S@naph,c3_c3@naph,c3_c2h1S@naph)|naphthalene

angle_coeff     7 class2/p6 bb    63.409397     1.435896     1.435896

angle_coeff     7 class2/p6 ba    33.560423    33.560423     1.435896     1.435896

angle_coeff     6 class2/p6      120.989674    30.577529   -24.527487     5.621286    -
4.025950     7.249626    # mm3->(c3_c3@naph,c3_c2h1S@naph,h1_c1S@naph)|naphthalene

angle_coeff     6 class2/p6 bb     7.682251     1.435896     1.090635

angle_coeff     6 class2/p6 ba    25.498143    28.515360     1.435896     1.090635

angle_coeff     3 class2/p6      116.807583    39.917331   -32.019324     7.338289    -
5.255662     9.463999    # mm3->(c3_c2h1@naph,c3_c2h1@naph,h1_c1@naph)|naphthalene

angle_coeff     3 class2/p6 bb     8.763731     1.454449     1.096263

angle_coeff     3 class2/p6 ba    30.724541    26.104046     1.454449     1.096263

angle_coeff     4 class2/p6      114.769310    93.785723   -75.229365    17.241300   -
12.348172    22.235656    # mm3->(c3_c2h1@naph,c3_c2h1S@naph,c3_c3@naph)|naphthalene

angle_coeff     4 class2/p6 bb    90.168492     1.391210     1.435896

angle_coeff     4 class2/p6 ba    86.536187    63.628816     1.391210     1.435896

angle_coeff     5 class2/p6      124.234272    29.012370   -23.272009     5.333552    -
3.819875     6.878542    # mm3->(c3_c2h1@naph,c3_c2h1S@naph,h1_c1S@naph)|naphthalene

angle_coeff     5 class2/p6 bb     9.732514     1.391210     1.090635

angle_coeff     5 class2/p6 ba    31.449745    26.623502     1.391210     1.090635

angle_coeff     1 class2/p6      112.413065    99.632083   -79.918970    18.316079   -
13.117925    23.621769    # mm3->(c3_c2h1@naph,c3_c2h1@naph,c3_c2h1S@naph)|naphthalene

angle_coeff     1 class2/p6 bb   104.912889     1.454449     1.391210
```

```
angle_coeff     1 class2/p6 ba   106.286485    77.649009     1.454449     1.391210

angle_coeff     2 class2/p6      118.001113    38.562917   -30.932893     7.089297    -
5.077335     9.142881    # mm3->(c3_c2h1S@naph,c3_c2h1@naph,h1_c1@naph)|naphthalene

angle_coeff     2 class2/p6 bb     9.862886     1.391210     1.096263

angle_coeff     2 class2/p6 ba    31.790211    26.090601     1.391210     1.096263


dihedral_style hybrid opls class2


dihedral_coeff     4 class2     0.000000     0.000000     2.279729     0.000000     0.000000
0.000000    # class2->(h1_c1@naph,c3_c2h1@naph,c3_c2h1S@naph,h1_c1S@naph)|naphthalene

dihedral_coeff     4 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff     4 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     4 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     4 class2 aat 0.0 0.0 0.0

dihedral_coeff     4 class2 bb13 0.0 0.0 0.0

dihedral_coeff    11 class2     0.000000     0.000000     2.415742     0.000000     0.000000
0.000000    # class2->(h1_c1S@naph,c3_c2h1S@naph,c3_c3@naph,c3_c3@naph)|naphthalene

dihedral_coeff    11 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff    11 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff    11 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff    11 class2 aat 0.0 0.0 0.0

dihedral_coeff    11 class2 bb13 0.0 0.0 0.0

dihedral_coeff     8 class2     0.000000     0.000000     1.897513     0.000000     0.000000
0.000000    # class2->(c3_c2h1@naph,c3_c2h1S@naph,c3_c3@naph,c3_c2h1S@naph)|naphthalene

dihedral_coeff     8 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff     8 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     8 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     8 class2 aat 0.0 0.0 0.0

dihedral_coeff     8 class2 bb13 0.0 0.0 0.0

dihedral_coeff     2 class2     0.000000     0.000000     3.996343     0.000000     0.000000
0.000000    # class2->(c3_c2h1@naph,c3_c2h1@naph,c3_c2h1S@naph,h1_c1S@naph)|naphthalene

dihedral_coeff     2 class2 mbt 0.0 0.0 0.0 0.0
```

```
dihedral_coeff     2 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     2 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     2 class2 aat 0.0 0.0 0.0

dihedral_coeff     2 class2 bb13 0.0 0.0 0.0

dihedral_coeff    10 class2     0.000000     0.000000     1.639432     0.000000     0.000000
0.000000    # class2->(h1_c1S@naph,c3_c2h1S@naph,c3_c3@naph,c3_c2h1S@naph)|naphthalene

dihedral_coeff    10 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff    10 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff    10 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff    10 class2 aat 0.0 0.0 0.0

dihedral_coeff    10 class2 bb13 0.0 0.0 0.0

dihedral_coeff    12 class2     0.000000     0.000000     2.799249     0.000000     0.000000
0.000000    # class2->(c3_c2h1S@naph,c3_c3@naph,c3_c3@naph,c3_c2h1S@naph)|naphthalene

dihedral_coeff    12 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff    12 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff    12 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff    12 class2 aat 0.0 0.0 0.0

dihedral_coeff    12 class2 bb13 0.0 0.0 0.0

dihedral_coeff     1 class2 bb13   -70.099189     1.454449     1.435896    # bb13-
>(c3_c2h1@naph,c3_c2h1@naph,c3_c2h1S@naph,c3_c3@naph)|naphthalene

dihedral_coeff     1 class2     0.000000     0.000000     3.801176     0.000000     0.000000
0.000000    # class2->(c3_c2h1@naph,c3_c2h1@naph,c3_c2h1S@naph,c3_c3@naph)|naphthalene

dihedral_coeff     1 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff     1 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     1 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     1 class2 aat 0.0 0.0 0.0

dihedral_coeff     3 class2     0.000000     0.000000     4.025101     0.000000     0.000000
0.000000    # class2->(h1_c1@naph,c3_c2h1@naph,c3_c2h1S@naph,c3_c3@naph)|naphthalene

dihedral_coeff     3 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff     3 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     3 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff     3 class2 aat 0.0 0.0 0.0
```

```
dihedral_coeff      3 class2 bb13 0.0 0.0 0.0

dihedral_coeff      9 class2      0.000000      0.000000      7.137224      0.000000      0.000000
0.000000    # class2->(c3_c2h1@naph,c3_c2h1S@naph,c3_c3@naph,c3_c3@naph)|naphthalene

dihedral_coeff      9 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff      9 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      9 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      9 class2 aat 0.0 0.0 0.0

dihedral_coeff      9 class2 bb13 0.0 0.0 0.0

dihedral_coeff      6 class2      0.000000      0.000000      2.585528      0.000000      0.000000
0.000000    # class2->(c3_c2h1S@naph,c3_c2h1@naph,c3_c2h1@naph,h1_c1@naph)|naphthalene

dihedral_coeff      6 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff      6 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      6 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      6 class2 aat 0.0 0.0 0.0

dihedral_coeff      6 class2 bb13 0.0 0.0 0.0

dihedral_coeff      5 class2 bb13   -75.085553      1.391210      1.391210     # bb13-
>(c3_c2h1S@naph,c3_c2h1@naph,c3_c2h1@naph,c3_c2h1S@naph)|naphthalene

dihedral_coeff      5 class2      0.000000      0.000000      6.412985      0.000000      0.000000
0.000000    # class2->(c3_c2h1S@naph,c3_c2h1@naph,c3_c2h1@naph,c3_c2h1S@naph)|naphthalene

dihedral_coeff      5 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff      5 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      5 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      5 class2 aat 0.0 0.0 0.0

dihedral_coeff      7 class2      0.000000      0.000000      1.336357      0.000000      0.000000
0.000000    # class2->(h1_c1@naph,c3_c2h1@naph,c3_c2h1@naph,h1_c1@naph)|naphthalene

dihedral_coeff      7 class2 mbt 0.0 0.0 0.0 0.0

dihedral_coeff      7 class2 ebt 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      7 class2 at  0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

dihedral_coeff      7 class2 aat 0.0 0.0 0.0

dihedral_coeff      7 class2 bb13 0.0 0.0 0.0


improper_style inversion/harmonic
```

```
improper_coeff      2      1.217508      0.000000     # harm-
>(c3_c2h1S@naph,c3_c2h1@naph,c3_c3@naph,h1_c1S@naph)|naphthalene

improper_coeff      1      5.430825      0.000000     # harm-
>(c3_c2h1@naph,c3_c2h1@naph,c3_c2h1S@naph,h1_c1@naph)|naphthalene

improper_coeff      3      7.389044      0.000000     # harm-
>(c3_c3@naph,c3_c2h1S@naph,c3_c2h1S@naph,c3_c3@naph)|naphthalene


special_bonds lj 0.00 0.00 1.00 coul 1.00 1.00 1.00
```

9. Super cell convergence

In order to converge the dynamical matrix, one must increase the size of the considered supercell until the obtained change in frequencies becomes negligible. Fig. S6 shows the supercell convergence behavior of the PBE/D3-BJ settings. While the 1×2×1 is still much too small, starting from a 2×2×2 supercell, all displayed bands already have the right dispersion. Only in the ΓA direction, the acoustic bands display slight differences compared to the bands of larger supercells. The chosen 2×3×2 supercell shows virtually no difference to the 3×3×3 supercell but comes at much lower cost. For this reason, the 2×3×2 supercell was considered for the reference DFT calculation. The reason why the chosen number of unit cells in the *b*-direction is larger (3 instead of 2) is to account for the shorter lattice constant in that direction (see main text). This results in a similar "probing radius" for the interatomic force constants in all three spatial directions, while keeping the total system size at a level small enough.

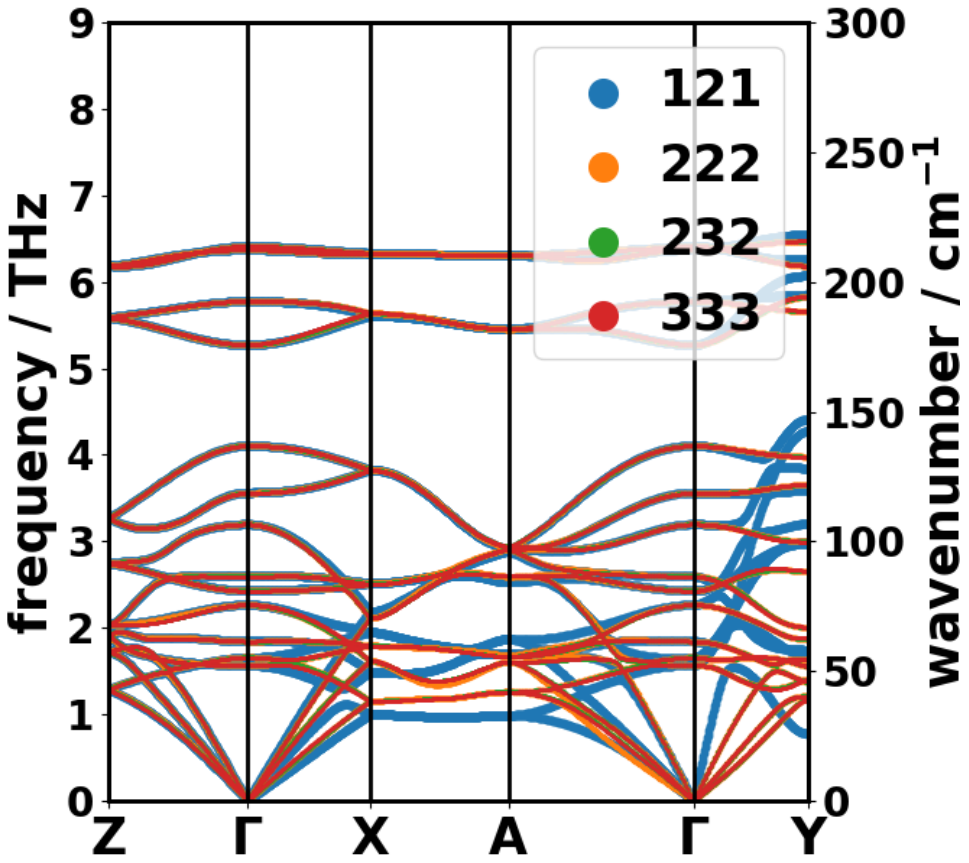

Fig. S6: *Convergence of phonon band structure (obtained with PBE/D3-BJ) with respect to supercell size.*

For the force field calculations, 3×3×3 supercells were used. As it is shown in Fig. S7 for the example of MOF-FF, this supercell is a good compromise between accuracy and effort, as there is no significant difference to a 4×4×4 supercell, while the 2×2×2 supercell shows minimal differences in ΓX direction.

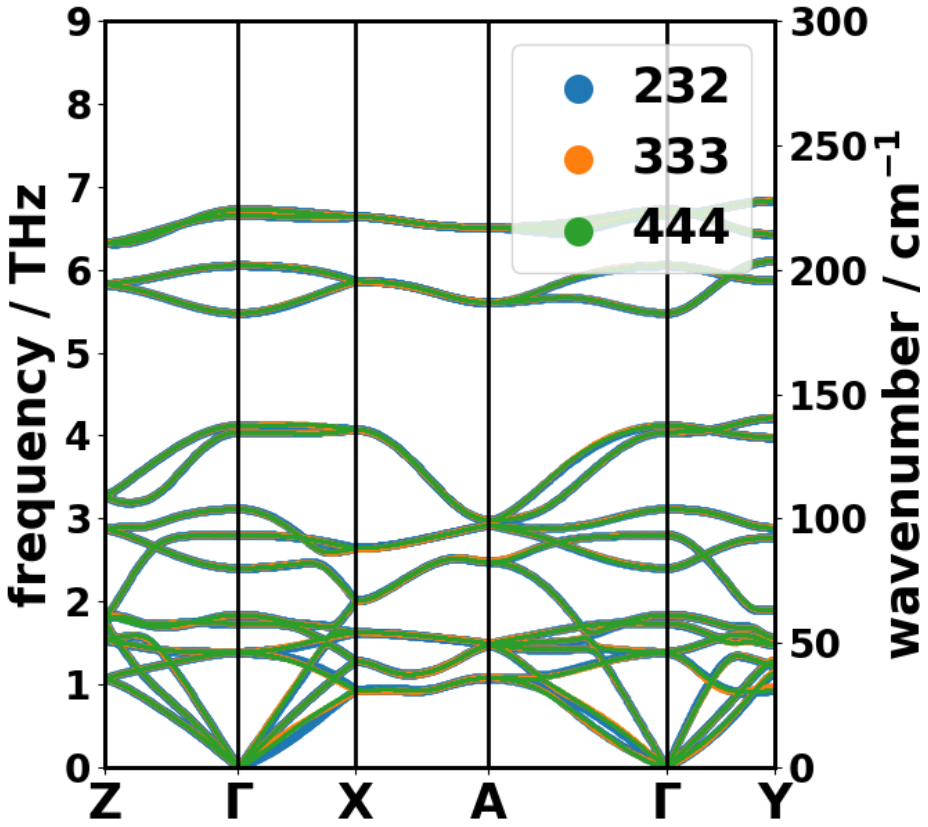

Fig. S7: *Convergence of phonon band structure (obtained with MOF-FF) with respect to supercell size.*

10. Comparison of Γ-frequencies for different van der Waals corrections

In the context of testing different van der Waals (vdW) corrections, the D3-BJ [8],[9], MBD [11],[12], D2 [13] and TS [14] schemes were used to fully optimize atomic coordinates and unit cells, and compute phonon frequencies. Since we have not been able to simulate phonon bands with the MBD correction, we compare the Γ-frequencies of the four approaches with the (few) experimental data points at this point in reciprocal space. This comparison in Fig. S8 shows that the MBD and the D3-BJ approach essentially yield the same phonon frequencies, while the other two approaches differ quite significantly in the low-frequency region (< 8 THz). As this frequency region is governed by intermolecular motion, the vdW interaction is especially important in this frequency regime.

At higher frequencies, the four approaches are in better agreement, although D2 tends to slightly underestimate frequencies above ~40 THz.

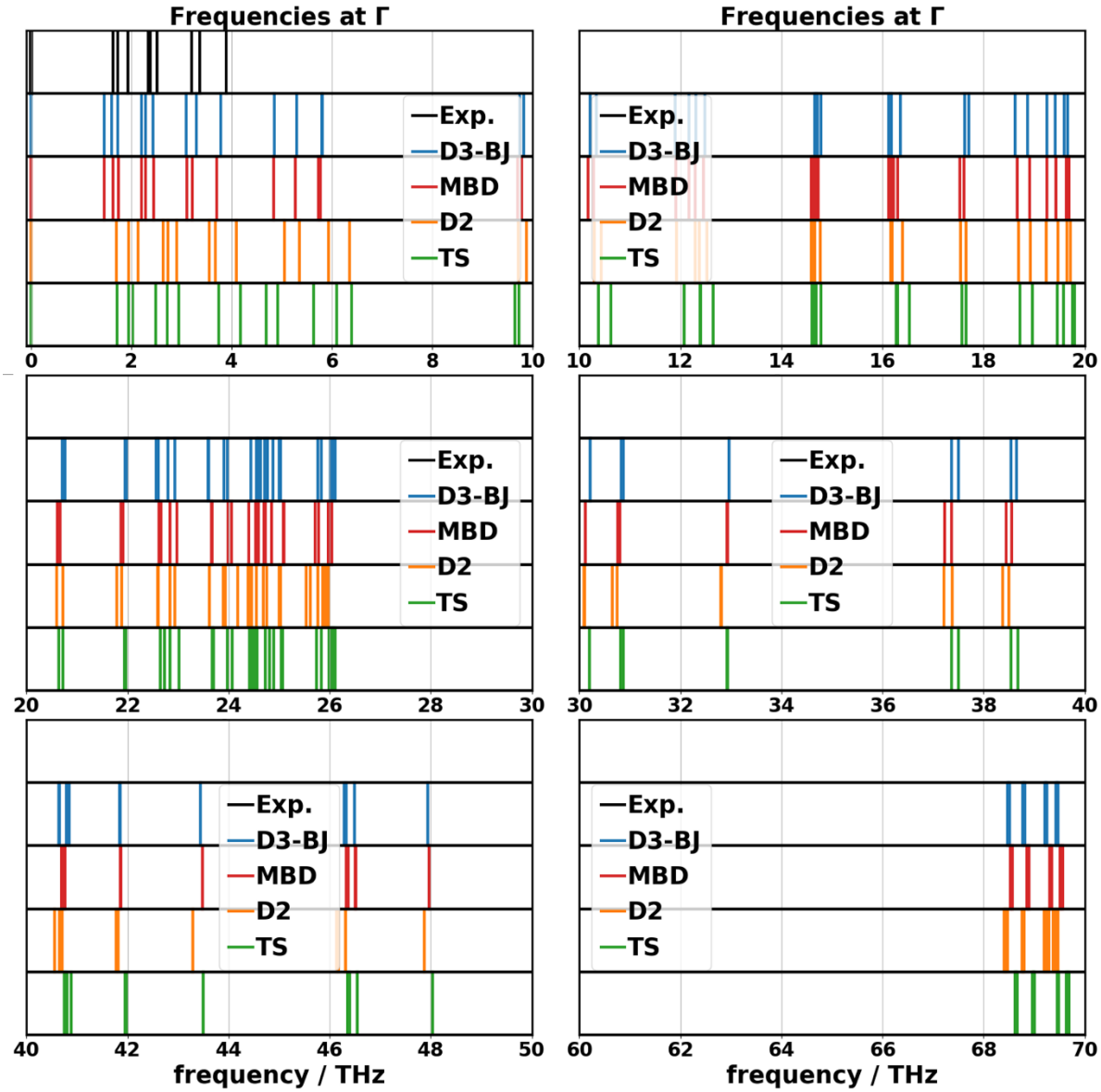

Fig. S8: *Comparison of Γ-frequencies obtained with different vdW corrections. The unit cells were fully relaxed employing the same correction.*

11. Comparison of phonon band structure of the deuterated systems to experiment

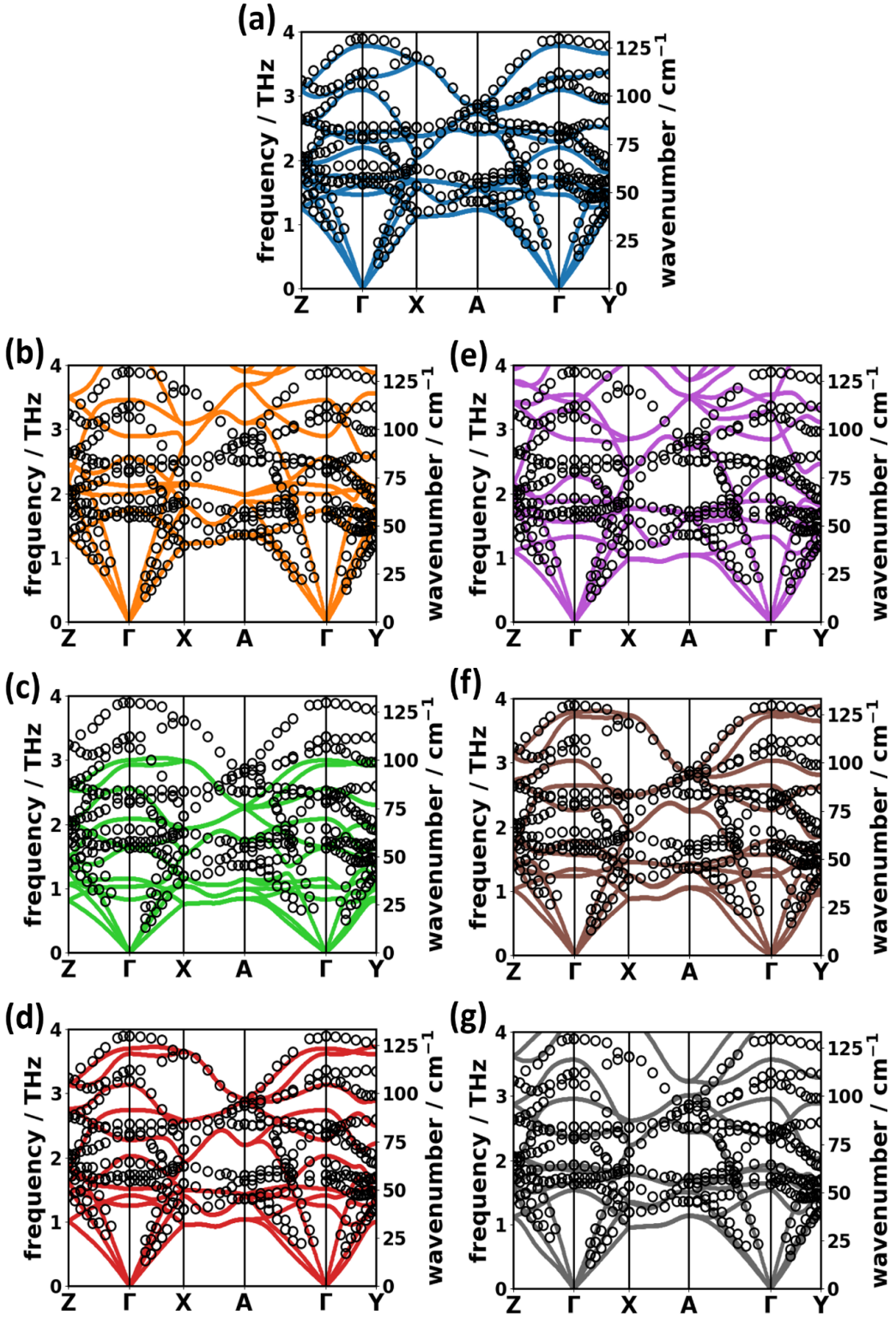

Fig. S9: *Phonon band structures (solid lines) of deuterated naphthalene obtained with (a) the DFT reference, (b) DFTB, (c) DFTB@DFT, (d) DFTB@95%DFT, (e) COMPASS, (f) MOF-FF, and (g) GAFF compared to experimental data points (open circles).*

As shown in Fig. S9, all trends discussed in the main manuscript for the comparison of approximate methods to DFT-D3BJ reference data are recovered also for the comparison to the experiments.

12. Assignment of vibrational modes

The so-called "Hungarian algorithm" of Kuhn [15] was used to solve the linear assignment problem – i.e. solving the problem of assigning each element of a set an element of a second set minimizing a cost function in that way. In our case, we have two sets of frequencies (phonons) stemming from two different methodologies, which we want to assign based on the similarity of their eigenvectors. In that problem the expression (3a) is minimized, with C_{ij} being a cost function and X being a matrix, whose entry X_{ij} is 1 if element i is assigned to element j. Here, X_{ij} is 1 if the i^{th} mode in the reference is assigned to the j^{th} mode of the comparison.

$$\sum_{ij} C_{ij} X_{ij} \tag{3a}$$

$$C_{ij} = \left(1 - S_{ij}\right) + P_{ij} \tag{3b}$$

$$S_{ij} = e^{\dagger}_{ref,i} \cdot e_{compared,j} \tag{3c}$$

$$P_{ij} = A\left(1 - \exp\left\{-\frac{\left(\omega_{ref,i} - \omega_{compared,j}\right)^2}{2\sigma^2}\right\}\right) \tag{3d}$$

The cost function must be adapted to the specific problem. Here, we define the cost function [see eq. (3b)] to consist of a matrix S_{ij} which characterizes the eigenvector overlap (complex dot product defined in eq. (3c); S_{ij} = 1 for perfect agreement, S_{ij} =0 for orthogonal eigenvectors) of the i^{th} eigenvector in the reference with the j^{th} eigenvector of the comparison. Additionally, we add a penalty function P_{ij} of Gaussian shape additionally penalizing mode assignments with large frequency differences [see eq. (3d)]. The parameters to choose are the amplitude A of the penalty and the Gaussian width σ. For the reported mode assignments, we used A=0.5 and σ=1 THz.

The penalty matrices, the eigenvector overlap and the resulting cost function for the mode assignment at the Γ-point are shown in Fig. S10. Note that especially COMPASS and GAFF tend

to show larger off-diagonal elements of the eigenvector overlap matrices, implying that the order of the eigenmodes changes compared to the DFT reference. Moreover, especially with these methods we observe several matrix entries significantly deviating from 1 and 0 indicating that the mode assignment is no longer unique. All tested methodologies show the biggest discrepancies in the nearly degenerate C-H stretching vibrations at ~90 THz.

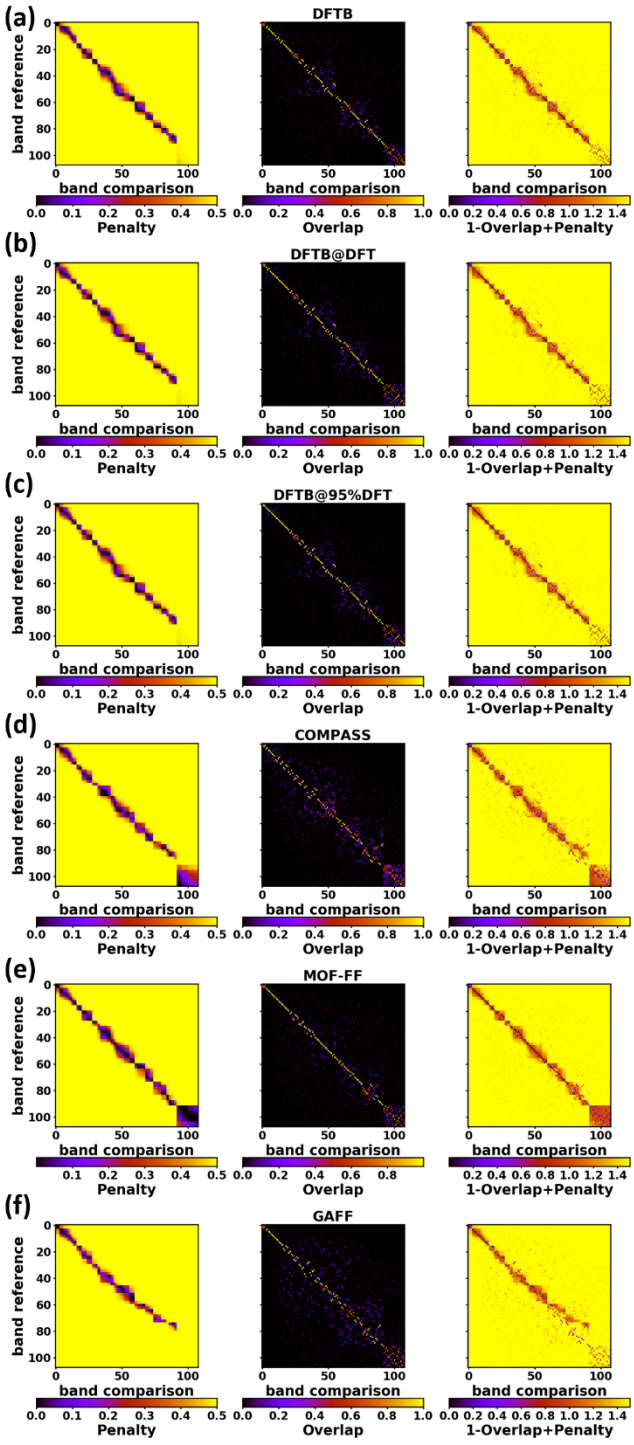

Fig. S10: *Penalty (left), overlap (center), and cost (right) functions of the (a) DFTB, (b) DFTB@DFT, (c) DFTB@95%, (d) COMPASS, (e) MOF-FF, and (f) GAFF phonon modes at Γ compared to DFT ref.*

13. Raman spectrum of molecular naphthalene using the B3LYP hybrid functional

In order to be consistent with the periodic DFT calculations, the molecular Raman spectra has first been calculated with the same functional. Additionally, the hybrid functional B3LYP [16],[17] has been used to obtain a comparison to a an approach more commonly applied when considering molecular systems [keeping the 6-311G(d,p)++ basis set and the D3-BJ van der Waals correction]. Fig. S11 compares the experimental data of Zhao and McCreery [18] with the crystal Raman spectrum (PBE/D3-BJ) and the molecular spectra (PBE/D3-BJ and B3LYP/D3-BJ). Notably, for the isolated molecule, the PBE simulation is in better agreement with the measured data than the spectrum calculated with the hybrid functional B3LYP, which slightly overestimates the vibrational frequencies. Moreover, the deviation between the B3LYP and PBE spectra is larger than the deviation between the PBE crystal and the PBE molecule simulation. It should be stressed that no empirical frequency scaling factors have been used.

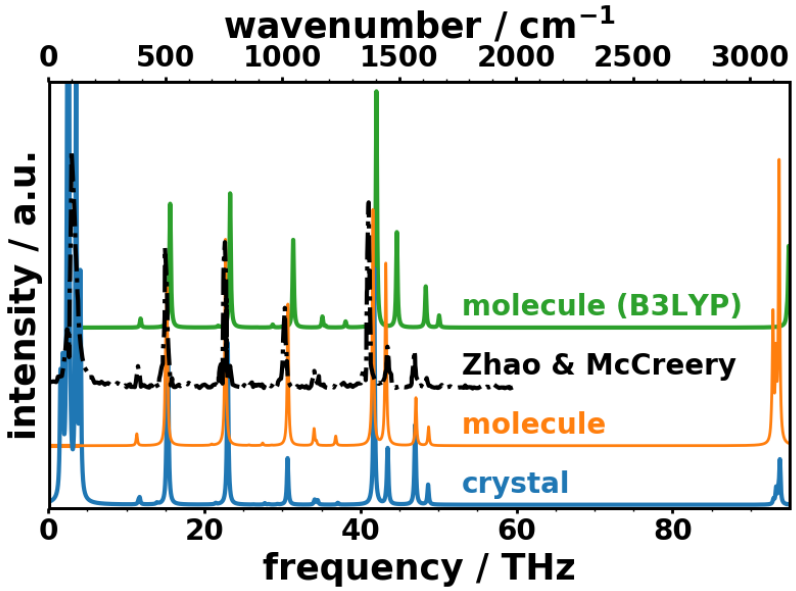

Fig. S11: *Simulated unpolarized Raman spectra for molecular (calculated with Gaussian16, 6-311G(d,p)++/PBE and 6-311G(d,p)/B3LYP) and crystalline (VASP, PBE) naphthalene (solid lines) compared to experimental data from Zhao and McCreery [18]. For both, the simulated spectra and the measurement of Zhao and McCreery, an excitation wavelength of 784 nm was used.*

14. Analysis of harmonic force constants

In order to track the origin of the most pronounced discrepancies in the phonon spectra, we will briefly comment on the relations of the observed (qualitative) discrepancies in the phonon band structures and the quantities that were actually calculated in the different approaches, namely the interatomic harmonic force constants (HFC). As the latter are rank 2 tensors, we base our discussion on the (rotational invariant) traces of the tensors to arrive at a qualitative, direction independent measure.

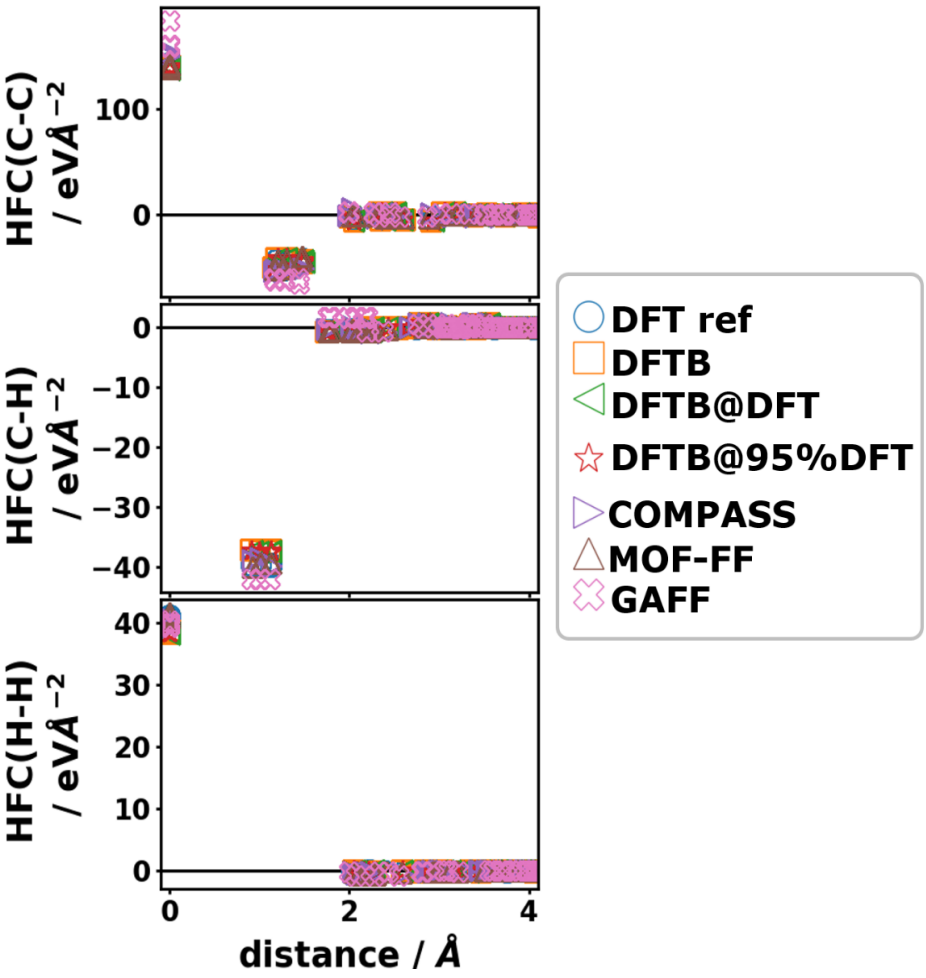

Fig. S12: *Trace of harmonic force constant tensors (HFCs) as a function of the distance between the involved atoms (considering periodic boundary conditions).*

The traces of the HFC tensors (in the following referred to as the HFCs) decrease very rapidly to zero (within ~3 Å) with the pair distance of the involved atoms (considering periodic boundary conditions) as shown in Fig. S12. This finding implies that all relevant interactions

can be seized within the spatial extents dictated by the considered supercell. Obviously, the HFCs are largest for such atom pairs which are covalently bonded (C-C and C-H) and are much greater in magnitude than HFCs between atoms of different molecules: the H-H HFCs are found to be smaller by two orders of magnitude compared to C-C or C-H HFCs.

The largest HFCs are found at zero pair distance and describe the interactions of an atom with itself (self-HFCs: SHFCs). SHFCs are typically the largest in magnitude because they correspond to the (negative) sum of interactions an atom is exposed to from all the other ones in the supercell, when it is displaced from its equilibrium position. It is important to note that SHFCs have a different sign than the HFCs between different atoms. These SHFCs are calculated by *PHONOPY's* internal routines by applying the translation invariance symmetries (*acoustic sum rules*). The SHFCs are convenient measures to assess the contributions from interactions beyond the nearest neighbor: if the SHFCs are much smaller than the (negative) sum of nearest neighbor HFCs, this suggests relevant long-range interactions.

It is, however, important to note that the cartesian HFCs can only be directly correlated to frequencies in simple cases, since the calculation of frequencies includes several mathematical operations like the diagonalization of the entire dynamical matrix. To be able to relate HFCs to directional atomic motion, one would have to analyze the HFC-tensors component-wise, which is beyond the scope of this work. Therefore, the HFCs cannot be directly used to draw quantitative conclusions about specific vibrations.

In the following we will briefly comment on the methodological trends before we discuss their impact on the observed phonon frequencies. A more detailed comparison of the HFCs obtained with the different methodology shows that there are relatively large differences in the individual values (see Fig. S13). It is shown in Fig. S13(a,b) that all considered DFTB-based approaches as well as MOF-FF yield comparable values of the C-C (S)HFCs, while, interestingly, the COMPASS FF and the GAFF show distinct discrepancies: (i) COMPASS overestimates the magnitudes of the C-C nearest neighbor interactions (i.e., they have large negative values for that force field). The fact that all nearest neighbor C-C HFCs in Fig. S13(c) are too large leads to the observed particularly large self-HFCs in Fig. S13(a). (ii) GAFF leads to an even more severe overestimation of the nearest neighbor C-C HFCs than COMPASS, giving rise to massively overestimated SHFCs.

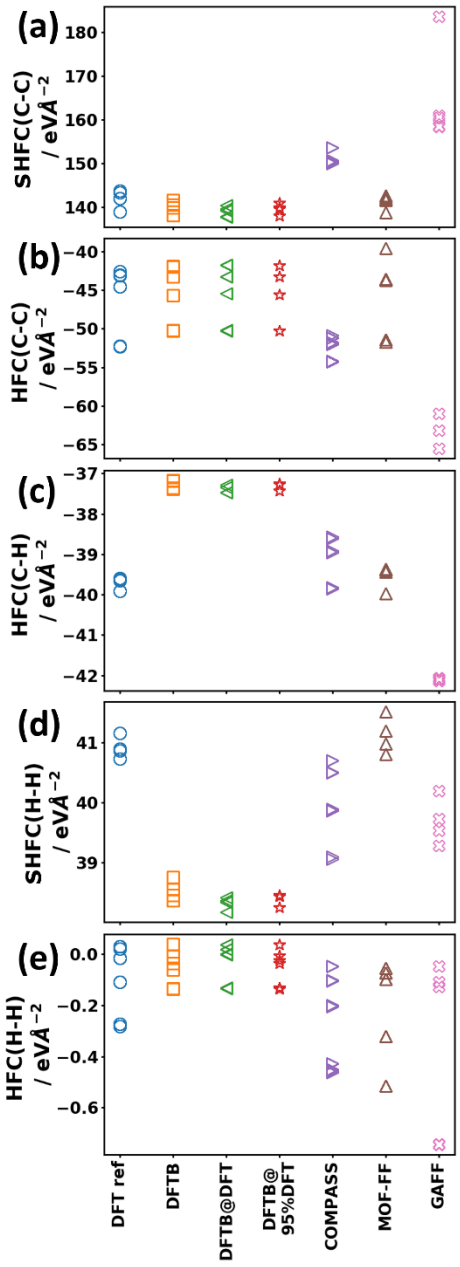

Fig. S13: *Harmonic force constants (HFCs) – i.e. traces of the rank 2 force constant tensors – obtained from different levels of theory. (a,d) Self-HFCs for carbon and hydrogen atoms, respectively. (b,c,e) HFCs for nearest neighbor interactions sorted according to the involved atomic species (C-C, C-H, H-H).*

Furthermore, Fig. S13(c-e) show that DFTB-based approaches consistently underestimate the magnitude of HFCs involving hydrogen, explaining the observation that the frequencies of C-H stretching vibrations are significantly underestimated. MOF-FF is able to reproduce the magnitude of the C-H HFCs. Relative the DFT reference data, GAFF yields discrepancies of similar magnitude as DFTB (albeit with a different sign). Interestingly, in all cases except for

the GAFF results, the H-H SHFCs approximately equal the negative sum of the shown nearest neighbor C-H [Fig. S13(c)] and H-H [Fig. S13(e)] interactions. GAFF, however, in both cases overestimates the respective HFCs, but still shows underestimated H-H SHFCs, which must be the (negative) sum of C-H and H-H HFCs. This finding is a strong indicator for relevant interaction beyond the pair distances of covalently bonded atoms. Indeed, HFCs can be found for the next-nearest C-H interactions (~1.8…2.3 Å distance), which differ by the reference HFCs at these distances by fa factor of ~-3.6 (see Fig. S14).

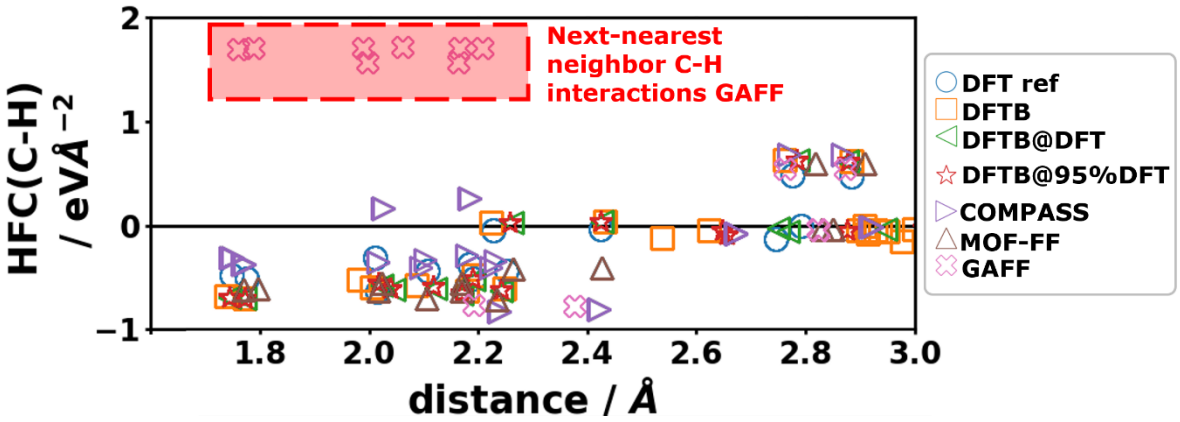

Fig. S14: *Trace of harmonic force constant tensors (HFCs) of C-H interactions as a function of distance of the involved atoms (considering periodic boundary conditions). The significantly overestimated HFCs from the GAFF are highlighted with a red box.*

The interpretation of the influence of the HFCs on the C-H stretching vibrations is slightly more involved. In the case of the DFTB-based approaches the interpretation is relatively straightforward: the too small C-H HFCs and H-H (S)HFCs are clear indicators why C-H stretching frequencies are underestimated that much. MOF-FF (COMPASS) shows slight tendencies to overestimate (underestimate) the corresponding frequencies, which is in agreement with the trends from the respective (S)HFCs. However, the GAFF results seem to be contradictive at first glance. The C-H HFCs are massively overestimated, while the H-H SFCs as well as the C-H stretching frequencies are too small. Therefore, in the case of GAFF, the influence of the underestimated H-H SFCs (due to the wrong next-nearest neighbor C-H HFCs; see Fig. S14) outweighs the effect of the overestimated C-H HFCs for the C-H stretching vibrations. The C-H bending frequencies are, however, overestimated by more than 10 THz. This observation is again in agreement with the too large C-H HFCs. This case is a good example that it is difficult and often impossible to draw direct conclusions from the HFCs for all but the simplest cases without a detailed analysis.

15. Differences in thermodynamic properties at higher temperatures

In the main manuscript, we show heat capacities up to temperatures of 400 K. At that temperature the heat capacity of naphthalene is still far from saturating at the classical limit. Because of the high-frequency C-H stretching vibrations (above ~90 THz), the envelope function f_C (see main text) must broaden to a large extent – i.e. by going to high temperatures – to reach those modes. For that reason, the heat capacity only approaches its classical limit ($3N\,k_B$) at temperatures beyond ~3500 K (see Fig. S15).

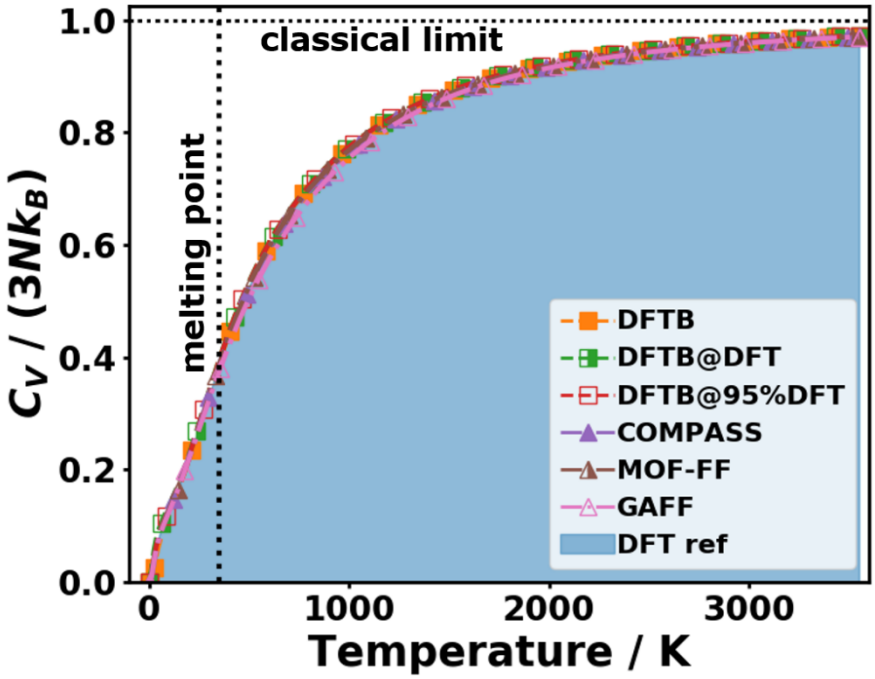

Fig. S15: *Saturation behavior of the phonon heat capacity C_V as a function of temperature.*

Although the melting point of crystalline naphthalene (at atmospheric pressure) is ~353 K [19], it is still instructive to hypothetically track the errors in thermal properties to higher temperatures. Fig. 16 shows the difference in (normalized) heat capacities with respect to the DFT/D3-BJ reference calculation. As the heat capacity is per definition a quantity that saturates at high temperatures (Dulong-Petit limit), the deviations also converge to zero. The fastest convergence is obviously reached by MOF-FF and COMPASS for the reasons discussed in the main text.

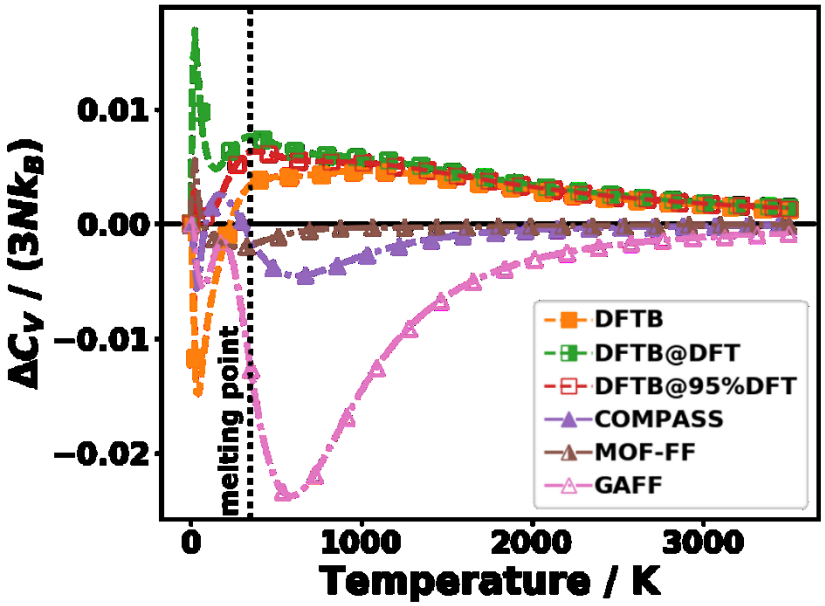

Fig. S16: *Difference in C_V with respect to the DFT reference (DFT ref) as a function of temperature shown for very large temperatures.*

The free energy, however, does not approach a saturation value. Therefore, the differences increase for large T (T $\rightarrow\infty$). Again for MOF-FF and COMPASS, the error at 3500 K is only ~±0.3 eV suggesting a relatively robust description of the free energy over a wide temperature range.

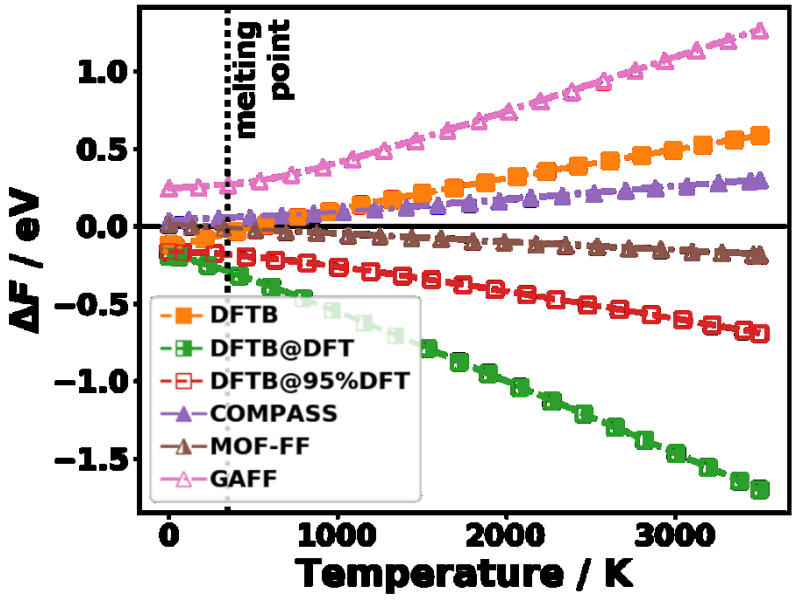

Fig. S17: *Difference in free energy F with respect to the reference as a function of temperature in a wide temperature range.*

REFERENCES

[1] H. Kuzmany, *Festkörperspektroskopie: Eine Einführung* (Springer, 1990).

[2] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J. V Ortiz, A.F. Izmaylov, J.L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M.J. Bearpark, J.J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, T.A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Cammi, J.W. Ochterski, R.L. Martin, K. Morokuma, O. Farkas, J.B. Foresman, and D.J. Fox, Gaussian 16 Revision 16.A.03 Inc., Wallingford CT, Inc., Wallingford CT (2016).

[3] G. Kresse, M. Marsman, and J. Furthmüller, VASP the Guide, (2018).

[4]DFTB+ 17.1 User Manual, (n.d.).

[5] S. Bureekaew, S. Amirjalayer, M. Tafipolsky, C. Spickermann, T.K. Roy, and R. Schmid, MOF-FF - A flexible first-principles derived force field for metal-organic frameworks, Phys. Status Solidi Basic Res. **250**, 1128 (2013).

[6] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, Electronic structure calculations on workstation computers: The program system turbomole, Chem. Phys. Lett. **162**, 165 (1989).

[7] J.P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. **77**, 3865 (1996).

[8] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. **132**, 154104 (2010).

[9] S. Grimme, S. Ehrlich, and L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, J. Comput. Chem. **32**, 1456 (2011).

[10] K. Eichkorn, O. Treutler, H. Ohm, M. Haser, and R. Ahlrichs, Auxiliary Basis-Sets to Approximate Coulomb Potentials (Vol 240, Pg 283, 1995), Chem. Phys. Lett. (1995).

[11] A. Tkatchenko, R.A. Distasio, R. Car, and M. Scheffler, Accurate and efficient method for many-

body van der Waals interactions, Phys. Rev. Lett. **108**, 1 (2012).

[12] A. Ambrosetti, A.M. Reilly, R.A. Distasio, and A. Tkatchenko, Long-range correlation energy calculated from coupled atomic response functions, J. Chem. Phys. **140**, (2014).

[13] S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion correction, J. Comput. Chem. **27**, 1787 (2006).

[14] A. Tkatchenko and M. Scheffler, Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data, Phys. Rev. Lett. **102**, 6 (2009).

[15] H.W. Kuhn, The Hungarian method for the assignment problem, Nav. Res. Logist. Q. **2**, 83 (1955).

[16] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. **98**, 5648 (1993).

[17] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, and M.J. Frisch, Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields, J. Phys. Chem. **98**, 11623 (1994).

[18] J. Zhao and R. Mccreery, Multichannel FT-Raman Spectroscopy: Noise Analysis and Performance Assessment, Appl. Spectrosc. - APPL SPECTROSC **51**, 1687 (1997).

[19] D.R. Lide, *CRC Handbook of Chemistry and Physics, 90th Edition* (Taylor & Francis, Boca Raton, Florida, 2009).