Robust Computation and Analysis of Vibrational Spectra of Layered Framework Materials including Host-Guest Interactions

Ekin Esme Bas,[*,†,‡,¶] Karen Marlenne Garcia Alvarez,[§] Andreas Schneemann,[§] Thomas Heine,[†,‡,¶,‖] and Dorothea Golze[*,†]

†*Chair of Theoretical Chemistry, Technische Universität Dresden, 01062 Dresden, Germany*
‡*Helmholtz-Zentrum Dresden-Rossendorf, HZDR, 01328 Dresden, Germany*
¶*Center for Advanced Systems Understanding, CASUS, 02826 Görlitz, Germany*
§*Chair of Inorganic Chemistry I, Technische Universität Dresden, 01069 Dresden, Germany*
‖*Department of Chemistry, Yonsei University and ibs-cnm, Seodaemun-gu, Seoul 120-749, Republic of Korea*

E-mail: ekin_esme.bas@tu-dresden.de; dorothea.golze@tu-dresden.de

Abstract

Layered framework materials, a rapidly advancing class of porous materials, are composed of molecular components stitched together via covalent bonds and are usually synthesized through wet-chemical methods. Computational infrared (IR) and Raman spectra are among the most important characterization tools for this materials class. Besides the *a priori* known spectra of the molecular building blocks and the solvent, they allow for *in situ* monitoring of the framework formation during synthesis. Therefore, they need to capture the additional peaks from host-guest interactions and the bands from emerging bonds between the molecular building blocks, verifying the successful synthesis of the desired material. In this work, we propose a robust computational framework based on *ab initio* molecular dynamics (AIMD), where we compute IR and Raman spectra from the time-correlation functions of dipole moments and polarizability tensors, respectively. As a case study, we apply our methodology to a covalent organic framework (COF) material, COF-1, and present its AIMD-computed IR and Raman spectra with and without 1,4-dioxane solvent molecules in its pores. To determine robust settings, we meticulously validate our model and explore how stacking disorder and different methods for computing dipole moments and polarizabilities affect IR and Raman intensities. Using our robust computational protocol, we achieve excellent agreement with experimental data. Furthermore, we illustrate how the computed spectra can be dissected into individual contributions from the solvent molecules, the molecular building blocks of COF-1, and the bonds connecting them.

1 Introduction

Two-dimensional (2D) covalent organic frameworks (COFs) are periodic layered materials composed of organic linkers connected by covalent bonds.[1,2] Due to high degree of porosity and large surface area, COFs are excellent candidates for a series of applications including gas storage[3,4] and gas separation.[5–8] Moreover, COFs have been reported to have promising application areas in sensing,[9,10] proton conduction,[11,12] catalysis,[13,14] energy storage,[15–18] light harvesting,[19,20] and electronics,[21–23] owing to a high amount of π-electron conjugation in-plane and π-π stacking interactions, which facilitate exciton migration and dissociation, that can be further enhanced via tuning the donor-acceptor units.[20]

Vibrational spectroscopy techniques, such as infrared (IR) and Raman spectroscopy, are widely used tools for the characterization of 2D materials, including identifying contributions from different components of host-guest systems.[24] Since 2D COFs,[1] and 2D polymers[25,26] are commonly syn-

thesized through wet-chemical methods, any remaining solvent contributes signals, potentially complicating the interpretation of the spectra. It has been previously reported for 2D COFs that the presence of guest molecules can affect several properties including the stacking order[27,28] and tensile strain.[29] A recent study reveals that pristine COF-1 and COF-1 with trapped solvent (mesitylene) molecules inside its pores show varying degrees of stability under high-pressure, which can be monitored via Raman spectroscopy.[29] Raman measurements are also employed to track reactions in solution. Some noteworthy examples include the *in situ* catalytic reaction monitoring of phthalocyanine-based conjugated COFs[30] and metal-organic frameworks (MOFs)[31] using *in situ* Raman spectroelectrochemistry.

The most common approach to compute vibrational spectra is normal mode analysis, typically within the Born-Oppenheimer approximation.[32] While this approach is computationally efficient, it has several limitations: It only allows the calculation of the fundamental transitions, which excludes a series of phenomena, such as overtones, combination and hot bands, Fermi resonances and bond breaking. Anharmonic effects are not captured and explicit solvent effects cannot be included because the structure must be in a local minimum and configurational sampling is thus not accounted for.

Ab initio molecular dynamics (AIMD) within the Born-Oppenheimer approximation is a promising alternative to overcome the limitations of the normal mode analysis, in particular for including anharmonicities along with temperature and solvent effects. However, this comes at the expense of increased computational cost in the case of systems with only a few hundreds of atoms, and the loss of nuclear quantum effects such as the vibrational zero-point energy or quantum tunneling[33,34] when treating nuclei classically.

AIMD as well as classical MD-based approaches provide an overall description of the modes activated at a specific temperature range. A large number of studies have been conducted on the computation of the vibrational spectra for various systems from dynamic approaches. The calculation of the vibrational spectrum of liquid water using classical MD and AIMD has been discussed in several studies,[35–38] along with other liquid,[39–41] gas-phase,[42–46] and air-water interface systems.[47–49] In the case of solids, the vibrational spectra of high-pressure ice,[50] hydrogen chloride hydrates,[51,52] crystalline naphthalene,[53] crystalline paracetamol,[54] zeolites,[55,56] metal-halide perovskites[56] and MOFs[56] have so far been investigated via MD-based approaches. The application of such methods for predicting IR and Raman spectra of 2D materials has been scarce. We are aware of two recent Raman studies on 2D YbOCl[57] and 2D titanium carbide MXene,[58] the latter one in combination with machine-learned force-fields.

In this work, we assess the AIMD-based approach for accurately predicting IR and Raman spectra of layered 2D materials. Our goal is to develop a computational tool to support the interpretation of *in situ* vibrational spectroscopy during wet-chemical synthesis of layered COFs and similar porous materials. We focus thus on host-guest interactions between the framework material and solvent molecules and strategies for analyzing the AIMD-based spectra. We demonstrate how to dissect the AIMD-based spectra into individual contributions from the solvent, the molecular units of the framework material, and the covalent bonds connecting these molecular units. Identifying and monitoring the vibrational bands related to these connecting bonds is crucial for verifying the successful synthesis of the framework material. Specifically, we investigate the COF-1 system with and without 1,4-dioxane solvent molecules in the pores. We propose computationally robust settings to obtain AIMD-based IR and Raman spectra for COF-1 and demonstrate how to dissect the spectra by adapting established strategies. We expect that the majority of our findings are transferable to other framework materials.

The structure of the paper is as follows: In Section 2, we provide an overview of the theory. Section 3 introduces the COF-1 test system and provides implementation details. Section 4 gives information on the computational procedure and the experimental setup. Section 5 includes details on the validation of our computational setup. We summarize and discuss our results for COF-1 in Section 6, comparing the MD-based results with the experimental data and the static calculations, analyzing the vibrational bands, and assessing host-guest interactions in the IR and Raman spectra. In Section 7, we provide our conclusions.

2 Theory

In the following, we summarize the key equations for the MD-based IR and Raman intensities,[59] together with the different methodologies we tested for the calculation of dipole moments and polarizability tensors. The expressions for IR and Raman intensities for static calculations based on conventional normal mode analysis are given in Section S1 in the Supplementary Information (SI).

2.1 IR and Raman spectra from MD

For MD-derived vibrational spectra, we generate a trajectory and calculate the dipole moments and polarizability tensors for a selected set of snapshots which are equally spaced over the trajectory. The vibrational spectra are obtained from autocorrelation functions (ACFs) containing these quantities.

The ACF of the real-valued property A is defined as:[60]

$$\langle A(\tau)A(\tau+t)\rangle_\tau = \int A(\tau)A(\tau+t)d\tau \quad (1)$$

$$\approx \frac{1}{\tau_{\max}}\sum_{\tau=1}^{\tau_{\max}} A(\tau)A(\tau+t) \quad (2)$$

In practice, we use Equation (2) since we have a set of discretized values for A. τ and t label the times for two different snapshots and $\tau_{\max}$ is the correlation depth, which is an input parameter.

The IR intensities are obtained from the Fourier transformations of the ACFs of $\dot{\boldsymbol{\mu}}$, which denotes the time derivative of the dipole moment $\boldsymbol{\mu}$.[61] The expression for I_{IR} is given by:

$$I_{\mathrm{IR}}(\tilde{\nu}) = \frac{2N_A}{12\varepsilon_0 c k_B T}\int_{-\infty}^{\infty}\langle\dot{\boldsymbol{\mu}}(\tau)\cdot\dot{\boldsymbol{\mu}}(\tau+t)\rangle_\tau \exp(-2\pi c\tilde{\nu}t)dt \quad (3)$$

where ε_0 is the vacuum permittivity, N_A is the Avogadro constant, c is the speed of light, k_B is the Boltzmann constant, T is the temperature and $\tilde{\nu}$ is the wavenumber.

Following Placzek's polarization theory, the total (unpolarized) Raman intensity is given by:[32,59]

$$I_{\mathrm{Raman}}(\tilde{\nu}) = \frac{2h}{8\epsilon_0^2 k_B T}\frac{(\tilde{\nu}_{in}-\tilde{\nu})^4}{\tilde{\nu}}\frac{1}{1-\exp\left(-\frac{hc\tilde{\nu}}{k_B T}\right)}\frac{45\delta^2(\tilde{\nu})+7\epsilon^2(\tilde{\nu})}{45} \quad (4)$$

where $\tilde{\nu}_{in}$ denotes the wavenumber of the incident light and h is the Planck's constant. The isotropic contribution $\delta^2(\tilde{\nu})$ and the anisotropic contribution $\epsilon^2(\tilde{\nu})$ are computed in the MD-based approaches from Fourier transforms of ACFs as follows:[62]

$$\delta^2(\tilde{\nu}) = \int_{-\infty}^{\infty}\left\langle\frac{\dot{\alpha}_{xx}(\tau)+\dot{\alpha}_{yy}(\tau)+\dot{\alpha}_{zz}(\tau)}{3}\times\frac{\dot{\alpha}_{xx}(\tau+t)+\dot{\alpha}_{yy}(\tau+t)+\dot{\alpha}_{zz}(\tau+t)}{3}\right\rangle_\tau \times \exp(-2\pi ic\tilde{\nu}t)dt \quad (5)$$

where $\dot{\boldsymbol{\alpha}}$ denotes the time derivative of the polarizability tensor $\boldsymbol{\alpha}$. The anisotropic contribution $\epsilon^2(\tilde{\nu})$ is given by:

$$\begin{aligned}\epsilon^2(\tilde{\nu}) = \int_{-\infty}^{\infty}\Big[&\frac{1}{2}\left\langle\left\{\dot{\alpha}_{xx}(\tau)-\dot{\alpha}_{yy}(\tau)\right\}\left\{\dot{\alpha}_{xx}(\tau+t)-\dot{\alpha}_{yy}(\tau+t)\right\}\right\rangle_\tau\\ &+\frac{1}{2}\left\langle\left\{\dot{\alpha}_{yy}(\tau)-\dot{\alpha}_{zz}(\tau)\right\}\left\{\dot{\alpha}_{yy}(\tau+t)-\dot{\alpha}_{zz}(\tau+t)\right\}\right\rangle_\tau\\ &+\frac{1}{2}\left\langle\left\{\dot{\alpha}_{zz}(\tau)-\dot{\alpha}_{xx}(\tau)\right\}\left\{\dot{\alpha}_{zz}(\tau+t)-\dot{\alpha}_{xx}(\tau+t)\right\}\right\rangle_\tau\\ &+\frac{3}{4}\left\langle\left\{\dot{\alpha}_{xy}(\tau)+\dot{\alpha}_{yx}(\tau)\right\}\left\{\dot{\alpha}_{xy}(\tau+t)+\dot{\alpha}_{yx}(\tau+t)\right\}\right\rangle_\tau\\ &+\frac{3}{4}\left\langle\left\{\dot{\alpha}_{yz}(\tau)+\dot{\alpha}_{zy}(\tau)\right\}\left\{\dot{\alpha}_{yz}(\tau+t)+\dot{\alpha}_{zy}(\tau+t)\right\}\right\rangle_\tau\\ &+\frac{3}{4}\left\langle\left\{\dot{\alpha}_{zx}(\tau)+\dot{\alpha}_{xz}(\tau)\right\}\left\{\dot{\alpha}_{zx}(\tau+t)+\dot{\alpha}_{xz}(\tau+t)\right\}\right\rangle_\tau\Big]\\ &\times\exp(-2\pi ic\tilde{\nu}t)dt\end{aligned} \quad (6)$$

The spectral width of the MD-based IR and Raman spectra is determined by the time interval Δt between the selected snapshots, where the maximum wavenumber is $\tilde{\nu}_{\max} = 1/(2c\Delta t)$. Choosing a time step of 2.5 fs corresponds to a spectral range up to a wavenumber of 6671.28 cm^{-1}. The resolution of the spectra is connected to the correlation depth $\tau_{\max}$ and can be enhanced by, e.g., data mirroring as described in Section 3.2.2.

The MD-based approach yields also access to the power spectrum, which is obtained from the ACFs of the velocities as shown in detail in Section S2 in the SI. The power spectrum displays all vibrational frequencies independently on IR and Raman selection rules and is the MD-equivalent to the set of normal frequencies obtained after diagonalizing the Hessian matrix in the static approach.

2.2 Calculation of the dipole moments

For finite systems, the electronic contribution to the dipole moment is given in terms of the expectation value of the position operator $\hat{\mathbf{r}}$ as $\boldsymbol{\mu} = -e\sum_i\langle\psi_i|\hat{\mathbf{r}}|\psi_i\rangle$, where ψ_i are the Kohn-Sham density functional theory (KS-DFT) orbitals. This definition does not generalize to periodic systems, where the KS-DFT orbitals are represented by Bloch functions $\psi_{n\mathbf{k}}$, where $\mathbf{k}$ is the wave vector. The expectation value $\langle\psi_{n\mathbf{k}}|\hat{\mathbf{r}}|\psi_{n\mathbf{k}}\rangle$ is ill-defined since the integrand diverges as $|\mathbf{r}|$ approaches infinity. The modern theory of polarization[63–65] solved this problem by providing a rigorous reformulation of the electric polarization (dipole moment per unit volume) in terms of a Berry phase or Wannier centers. In the following, we summarize the key equations and refer the reader to Ref. 66 for a detailed description.

2.2.1 Berry phase approach

The modern theory of polarization recasts the expression for the expectation value of the position operator into a Berry phase of the cell-periodic parts of the Bloch functions.[63–65] This approach abandons the notion of a completely unique value for the dipole moment, recognizing instead that the

latter is determined modulo a so-called quantum of polarization. In this work, we employ large supercells for the computation of IR and Raman intensities. For large supercells, an evaluation only at the Γ-point is justified, for which the Berry phase dipole expression reads:

$$\mu_\alpha^{\text{Berry}} = \frac{e}{2\pi}\sum_\beta h_{\alpha\beta}\gamma_\alpha + \mu_\alpha^{\text{nuc}} \tag{7}$$

where α and β indicate the Cartesian directions (x,y,z). $h_{\alpha\beta}$ is an element of the 3×3 matrix $\mathbf{h} = [\mathbf{a},\mathbf{b},\mathbf{c}]$, which defines the simulation cell with the primitive Bravais lattice vectors $\mathbf{a},\mathbf{b}$ and $\mathbf{c}$. μ_α^{nuc} is the nuclear contribution. The α component of the Berry phase $\boldsymbol{\gamma}$ is given as:

$$\gamma_\alpha = \text{Im}\ln\det \mathbf{S}^{(\alpha)} \tag{8}$$

with

$$S_{nm}^{(\alpha)} = \langle\psi_n \,|\, \exp\{-i\mathbf{G}_\alpha\cdot\mathbf{r}\} \,|\, \psi_m\rangle \tag{9}$$

and the reciprocal lattice vectors:

$$\mathbf{G}_\alpha = 2\pi(\mathbf{h}^{-1})^{\text{T}}\mathbf{e}_\alpha \tag{10}$$

with $\mathbf{e}_x = (1,0,0)$, $\mathbf{e}_y = (0,1,0)$ $\mathbf{e}_z = (0,0,1)$. $\psi_{n/m}$ are *occupied* KS-DFT orbitals under periodic boundary conditions (PBCs).

The Berry phase approach yields the total electric dipole moment for the whole system. Since $\exp\{i\mathbf{G}_\alpha\cdot\hat{\mathbf{r}}\}$ in Equation(9) is a periodic function, the Berry phase dipole moment is only defined modulo the lattice vector. During an MD simulation, $\mu_\alpha^{\text{Berry}}$ can "jump", which needs to be addressed by adding or subtracting multiples of $a_\alpha + b_\alpha + c_\alpha$ to or from the α component of the dipole vector, to ensure that we autocorrelate dipole moments from the same polarization branch.

2.2.2 Maximally localized Wannier functions

Wannier functions are derived through a unitary transformation of the Bloch functions to localized functions. Specifically, we employ maximally localized Wannier functions (MLWFs),[67,68] obtained by applying a "maximal-localization" criterion to produce a unique set of Wannier functions. The minimization of the spread of the MLWFs is based on the concepts of the modern theory of polarization. While the regular position operator can be used for finite systems, the spread function is derived in terms of a Berry phase under PBCs. MLWFs are localized in space around the Wannier centers $\mathbf{r}_i^{\text{WC}}$. Adopting again a Γ-only formulation, the α component of the Wannier center is given by[68,69]

$$r_{\alpha i}^{\text{WC}} = -\frac{1}{2\pi}\sum_\beta h_{\alpha\beta}\text{Im}\ln\langle w_i \,|\, \exp\{-i\mathbf{G}_\alpha\cdot\mathbf{r}\} \,|\, w_i\rangle \tag{11}$$

where w_i is the MLWF and $\mathbf{G}_\alpha$ is defined in Equation (10).

Assuming a spin-unpolarized calculation, each MLWF is occupied by two electrons. Or in other words, we have a charge of $-2e$ at each Wannier center $\mathbf{r}_i^{\text{WC}}$. The electric dipole moment, including nuclear contributions, can then be evaluated as:

$$\boldsymbol{\mu}^{\text{MLWF}} = -2e\sum_i \tilde{\mathbf{r}}_i^{\text{WC}} + e\sum_i Z_i\tilde{\mathbf{R}}_i \tag{12}$$

where the distance vectors $\tilde{\mathbf{r}}_i^{\text{WC}} = \mathbf{r}_i^{\text{WC}} - \mathbf{r}_0$ and $\tilde{\mathbf{R}}_i = \mathbf{R}_i - \mathbf{r}_0$ are implemented with PBCs and $\mathbf{r}_0$ is a reference point, for example, the center of mass. $\mathbf{R}_i$ and Z_i indicate the atomic positions and charges, respectively. Equation (12) yields the total dipole moments of the supercell. As for the Berry phase dipole moments, "jumps" of $\boldsymbol{\mu}^{\text{MLWF}}$ during the AIMD are possible, which are again corrected by adding/subtracting multiples of $a_\alpha + b_\alpha + c_\alpha$.

A molecular dipole moment μ^{MLWF} is obtained by restricting the sums in Equation (12) to the Wannier centers and atomic charges assigned to the individual molecule. The sum of the molecular dipoles corresponds then to the total dipole moment of the cell.[39] In the context of IR and Raman, a typical application of MLWFs is to compute subspectra of liquid mixtures,[70] summing up only the molecular dipole moments of a subset of the molecules. Similarly, we employ MLWFs in this work to dissect the obtained spectra into contributions from solvent molecules (guest) and the material (host) itself, and also to analyse the spectral bands via breaking down the material into molecular units and connecting bonds. The definition of a "Wannier molecule", to which the $\mathbf{r}_i^{\text{WC}}$ can be uniquely assigned, is only approximately possible for framework materials. We elaborate our approach in Section 3.2.1.

2.3 Calculation of the polarizabilities

We tested two different approaches to compute polarizabilities. The first one is based on an induced dipole approach, whereas the second approach relies on a direct calculation of the polarizabilities within the framework of the density functional perturbation theory (DFPT).

2.3.1 Polarizabilities via induced dipoles

Polarizabilities can be calculated numerically through induced dipole moments $\boldsymbol{\mu}_{\text{ind}}$,[39] requiring no further modification to the electronic structure beyond the computation of electric dipole moments. The polarizability tensor $\boldsymbol{\alpha}$ is related to the change in the dipole moment under the applied external electric field $\mathbf{E} = (E_x, E_y, E_z)$ by:

$$\boldsymbol{\mu}_{\text{ind}} = \boldsymbol{\alpha}\mathbf{E} \tag{13}$$

In practice, we perform three additional single-point calculations for each of the selected MD snapshots. In these calculations, a finite periodic electric field E, which is computed using the Berry phase approach[71,72] and has a field

strength of 0.005 a.u., is applied along the x, y or z direction, i.e., we apply the fields $(E,0,0)$, $(0,E,0)$ or $(0,0,E)$. After convergence of the electronic structure, we compute the electric dipole moment $\boldsymbol{\mu}^k$, where k indicates the direction in which the field has been applied. The induced dipole moment is then calculated via forward finite differences as:

$$\boldsymbol{\mu}^k_{\text{ind}} = \boldsymbol{\mu}^k - \boldsymbol{\mu}^0 \quad \text{with} \quad k = x,y,z \tag{14}$$

where $\boldsymbol{\mu}^0$ denotes the dipole moment vector obtained from the field-free calculations. The components $\alpha_{\alpha\beta}$ of the polarizability tensor are then given by $\alpha_{\alpha,\beta=k} = \mu^k_{\text{ind},\alpha}/E$.

The dipole moments $\boldsymbol{\mu}_k$ and $\boldsymbol{\mu}_0$ can be calculated using different approaches, e.g., the Berry phase approach or MLWFs. The induced dipole moment is unique since we take the difference between two dipole moments on the same polarization branch. The advantage compared to the DFPT-based approach (Section 2.3.2) is that IR and Raman spectra can be treated on equal footing, meaning that the dipole moment (IR) and the induced dipole moment/polarizability (Raman) can be computed with the same method. This also enables the computation of local contributions in Raman spectra by computing the induced dipole moments via MLWFs. A disadvantage of the induced dipole strategy is that the field strength E has to be selected with care such that i) the finite difference approach is valid and ii) that non-linear effects, which might be relevant for larger E-fields, can be neglected. However, employing field strengths that are too low can introduce noise into the spectrum, necessitating a trial-and-error procedure to identify the optimal field strength for each material.

2.3.2 Polarizabilities via DFPT

DFPT[73,74] is an analytical approach to compute the response of a molecule or material to an external electric (or magnetic) field. It can be viewed as the solid-state formulation of the coupled perturbed self-consistent field (CPSCF) method[75,76] originally developed in the quantum chemistry community. The derivation of the DFPT equations relies on perturbation theory. The KS Hamiltonian is modified by an additional term $\hat{h}_{\text{E}} = -\hat{\mathbf{r}} \cdot \mathbf{E}$, where $\mathbf{E}$ is the applied homogeneous electric field. $\mathcal{E}_E$ denotes the energy term due to $\hat{h}_{\text{E}}$.

Expanding the total energy in a perturbative manner under the zero-field condition results in:

$$\mathcal{E}_{\text{tot}}(\mathbf{E}) \approx \mathcal{E}^0_{\text{tot}}[n_0] + \sum_\alpha \mu_\alpha E_\alpha + \frac{1}{2}\sum_{\alpha,\beta} \alpha_{\alpha\beta} E_\alpha E_\beta + \ldots \tag{15}$$

where $\mathcal{E}^0_{\text{tot}}[n_0]$ is the total energy of the unperturbed KS system with the ground-state density n_0 and α,β indicate again the Cartesian directions (x,y,z). The second-order term corresponds to the polarizability and is expressed within PBCs for the Γ-only case as:[77]

$$\alpha_{\alpha\beta} = \left.\frac{\partial^2 \mathcal{E}_{\text{E}}[n_0]}{\partial E_\alpha \partial E_\beta}\right|_{\mathbf{E}=0} = \left.\frac{\partial \mu^{\text{Berry}}_\alpha}{\partial E_\beta}\right|_{\mathbf{E}=0} \tag{16}$$

$$= \frac{e}{2\pi} \sum_\beta h_{\alpha\beta} \text{Im}\left[\sum_{nm} \left(\left\langle \psi^{(1)\beta}_n \middle| \exp[-i\mathbf{G}_\alpha \cdot \mathbf{r}] \middle| \psi^{(0)}_m \right\rangle + \left\langle \psi^{(0)}_n \middle| \exp[-i\mathbf{G}_\alpha \cdot \mathbf{r}] \middle| \psi^{(1)\beta}_m \right\rangle\right)\left(S^{(\alpha)}_{mn}\right)^{-1}\right] \tag{17}$$

where $\psi^{(0)}_n$ and $\psi^{(1)\beta}_n$ are the unperturbed and first-order perturbed KS-DFT orbitals and $S^{(\alpha)}_{nm}$ is defined in Equation (9). The first-order perturbed KS-DFT orbitals are obtained via a linear-response calculation. For more details of the DFPT procedure, see Ref. 44.

The DFPT polarizabilities are obtained for the whole supercell, while individual contributions are not accessible, unless modifications are made to the DFPT code.[40] An important advantage compared to the induced dipole method introduced in Section 2.3.1 is that the polarizabilities are directly obtained after a single linear-response calculation, requiring no manual adjustments of field strengths or post-processing steps. We note that also the DFPT polarizabilities are uniquely defined due to multiplication with the inversion of $S^{(\alpha)}_{nm}$ in Equation (17), canceling the phase indeterminancy.

3 System and Implementation

3.1 Test system COF-1

We selected COF-1 as our test system since it is well-characterized in terms of structural properties and vibrational modes.[28,78] COF-1 is a crystalline material composed of atomically thin 2D periodic layers. The COF-1 monolayer is a porous network consisting of connected phenyl and boroxine rings. The layered bulk material is formed by stacking the 2D layers via non-bonded interactions.

The stacking structure of COF-1 and other 2D COFs is not eclipsed as assumed from early low-resolution powder X-ray diffraction data (PXRD),[1] but between two layers there is a small shift of about 1.6 Å along one of the high-symmetry directions. This stacking is denoted as AA' and shown in Figure 1a.[79,80] These apparently small shifts can impose significant changes in the electronic properties.[81] In realistic COFs, there is no pronounced interaction between second next layers, yielding to a stochastic sequence of small shifts, which result in non-periodic structures normal to the COF plane.[28]

Accommodating a statistical stacking model is computationally too expensive for AIMD simulations. We employed thus COF-1 models with AA' stacking for the static calculations and as starting structures for the AIMD simulations. However, we tried to take the statistical stacking model somewhat into account by repeating the unit cell shown in

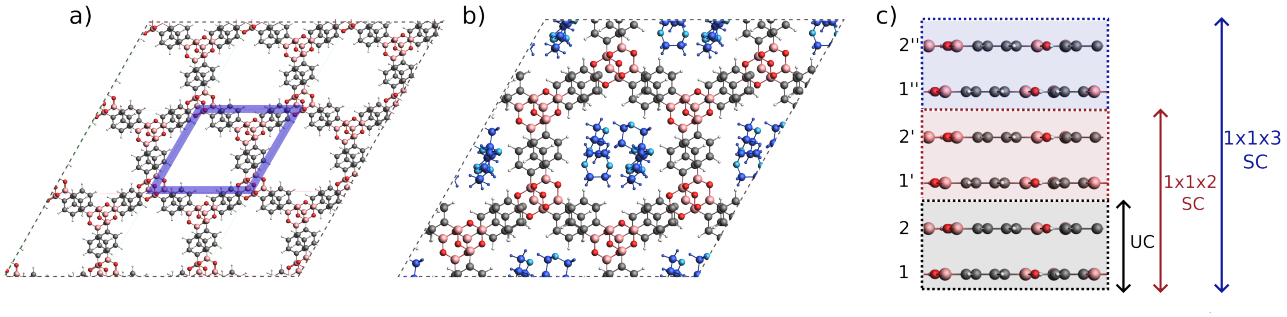

Figure 1 a) COF-1 system with AA' stacking. The unit cell is indicated by a purple rhombus. b) COF-1 system with the solvent molecules (1,4-dioxane) inside the pores. Color code: black: C, white: H, red: O, pink: B. The solvent molecules are depicted in blue. A lighter blue shade is used for the oxygen atoms for clarity. c) Side view representation showing the number of layers of the COF-1 unit cell, $1\times1\times2$ supercell used for the pristine COF-1, and the $1\times1\times3$ supercell used for the COF-1+solvent system.

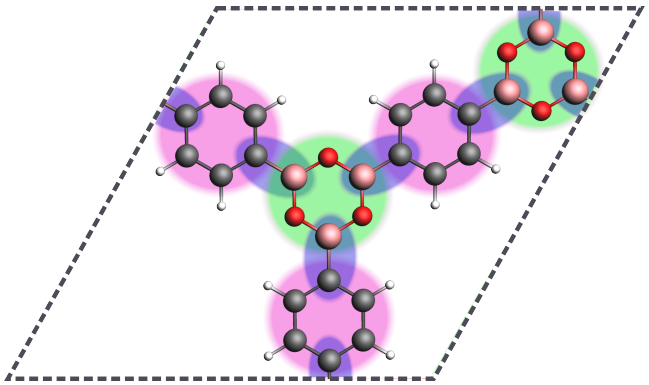

Figure 2 Molecular structure of COF-1 layer (unit cell) and the assigned fragments. The boroxine and phenyl fragments are highlighted in green and purple, the B-C bonds are highlighted in blue.

Figure 1(c) 2-3 times via a supercell approach, which is described in detail in Section 4.1. In the supercell model, symmetry is not enforced normal to the plane of the layers during the AIMD simulation, allowing as much independent stacking shifts as the selected supercell can afford. In a $1\times1\times2$ supercell, layers 1 and 1' will not remain aligned on top of each other during the AIMD simulation, see Figure 1c for labels.

For the investigation of solvent effects in the IR and Raman spectra, we included 1,4-dioxane molecules in the pores of our COF-1 model, as shown in Figure 1b. 1,4-dioxane is one of the solvents commonly used during the synthesis of COF-1.[1]

3.2 Implementation details

3.2.1 Fragment approach for MLWFs

The AIMD-based scheme does not give access to the individual vibrational modes, but only the total spectrum. Different methods were proposed to dissect the MD-based spectra into individual contributions, ranging from Voronoi integration techniques[82] to subsystem DFT-based embedding,[70] spatial decomposition methods[83] to extracting effective normal modes from the vibrational density of states.[84] The most common approach is decomposition via MLWFs,[67,85,86] which, from the practitioner's perspective, is straightforward to apply without requiring any modifications to the electronic-structure code. In the case of, e.g., molecular liquids, MLWFs can be easily assigned to the individual molecules. Such an assignment is less obvious for materials. However, COF frameworks are composed of "molecular units" which can be considered as separate fragments even though they are connected with covalent bonds.

We implemented such a fragment-based approach for COF-1, dividing the network in phenyl and boroxine fragments and boron-carbon (B-C) bonds as displayed in Figure 2. We computed and summed up the individual dipole moments for each of these fragments and B-C bonds to obtain their subspectra.

3.2.2 Computation of IR and Raman spectra

We evaluated the time derivatives of $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$, which are required in the ACFs (Equations (3) to (6)), through central finite differences:

$$\dot{x} = \frac{\delta x(t)}{\delta t} \approx \frac{x(t+\Delta t) - x(t-\Delta t)}{2\Delta t} \tag{18}$$

where Δt indicates the time step between the selected equidistant MD snapshots and x is a component of $\boldsymbol{\mu}$ or $\boldsymbol{\alpha}$. In addition, we applied a finite difference correction by dividing the obtained frequencies $I_{\mathrm{IR}}(\tilde{\nu})$ and $I_{\mathrm{Raman}}(\tilde{\nu})$ by the sinc function $\left(\frac{\sin(\tilde{\nu}\Delta t)}{\tilde{\nu}t}\right)^2$.[59,87] This correction factor alters the amplitude of the vibrations without affecting the frequencies. We carried out the discrete Fourier transformations of the ACFs of $\dot{\boldsymbol{\mu}}$ or $\dot{\boldsymbol{\alpha}}$ employing an interface to the FFTW 3.3 library.[88]

Prior to conducting the Fourier transformation, we applied so-called "data mirroring" to the ACFs.[59] That means, we added the data set in reverse order to the original autocorrelation data. Data mirroring does not change the integral of the spectrum, however it is useful to increase the resolution of the spectrum since it doubles the data points. To further enhance the resolution of the spectra, we employed window functions. More specifically, we multiplied the autocorrelation data with the Hann window function $\cos^2[\pi/(2\tau_{\max}-2)]$ before performing the Fourier transformations.[59] $\tau_{\max}$ is the correlation depth defined in Equation (2).

For the validation of the methodology and the results, we compared the computed Power, IR and Raman spectra with the spectra generated with the TRAVIS program package,[39,89,90] which targets application to molecular liquids. The spectra presented for liquid water AIMD simulations (Figure S4, SI) show perfect agreement with our results.

4 Methodology

4.1 Computational details

We carried out AIMD simulations within the Born-Oppenheimer approximation using the FHI-vibes package[91] interfaced to the FHI-aims code[92] for the computation of energies and forces. FHI-aims is an all-electron code based on numerical atom-centered orbitals (NAOs). The simulations were performed at 300 K with a time step of 0.5 fs, using the NVT ensemble with a Langevin thermostat.[93] The electronic structure calculations were performed with the Perdew–Burke-Ernzerhof (PBE) exchange-correlation functional[94] in combination with the Tkatchenko-Scheffler (TS) correction[95] to account for long-range dispersion interactions. NAOs of *tier1* quality in combination with "light" numerical settings were used.[92] The AIMD simulations were equilibrated for approximately 5.0 ps followed by a production run of at least 25.0 ps.

For the pristine COF-1 material (no solvent), we generated AIMD trajectories with two different setups using either *k*-point (KP) meshes or large supercells (SC). In the KP setup, our computational model is a COF-1 unit cell, i.e., a structure with 84 atoms shown in Figure 1c, in combination with a $1\times1\times3$ *k*-point grid. For the SC setup, we used a $1\times1\times2$ COF-1 supercell, which contains 4 layers and 168 atoms (see Figure 1c), and performed all electronic-structure calculations at the Γ-point only. In order to investigate the solvent effects, we generated a third AIMD trajectory using a $1\times1\times3$ COF-1 supercell with 1,4-dioxane molecules in the pores, see Figures 1b and c. The initial COF-1+solvent structure was generated with the PACKMOL package,[96] setting a density of 1.033 g/mL for the solvent. The COF-1+solvent supercell contains seven 1,4-dioxane molecules in the pores, resulting in a system size of 350 atoms. The AIMD starting structures were cell- and geometry optimized at the PBE+TS level with FHI-aims.

We extracted at least 5000 snapshots from the AIMD trajectories, evenly spaced by 2.5 fs. For each snapshot, we computed MLWFs, Berry phase dipole moments and polarizabilities using the CP2K program package.[97] For all CP2K calculations, we employed the PBE exchange-correlation functional,[94] Goedecker–Teter–Hutter (GTH) pseudopotentials[98–100] and the DZVP-MOLOPT-SR-GTH basis sets.[101] The CP2K calculations of the dipole moments and polarizabilities were conducted only at the Γ−point. Snapshots from the AIMD trajectories with the SC setup were used without modification, while the snapshots from the KP trajectories were extended to an $1\times1\times2$ supercell.

We calculated the polarizabilities via induced dipole moments and alternatively via DFPT. For the induced dipole method, we computed the MLWFs and Berry phase dipole moment under the influence of an external periodic electric field with an absolute value of 0.005 a.u. in x, y or z direction, respectively. We computed the ACFs of the dipole moments and polarizabilities with an autocorrelation depth of 10.24 ps and set an incident laser wavelength of 1064 nm for the computation of the Raman intensities. We report the unpolarized (total) Raman intensities in alignment with the experimental measurements.

We used FHI-aims instead of CP2K for the force evaluation since the TS dispersion correction, added as post-processing step to the electronic forces, is not implemented in CP2K. A proper choice of the dispersion correction model is crucial to keep the AIMD trajectories stable, as discussed in detail in Section 5.1. However, the dispersion correction is irrelevant for the electronic-structure optimization and does consequently not affect the computation of the dipole moments and polarizabilities for the single snapshots. Figure S5 (SI) demonstrates that the CP2K and FHI-aims calculations generate very similar electronic potential energy surface (PES). This ensures that the frequencies (via forces/velocities from FHI-aims) and the corresponding IR/Raman intensities (from CP2K) are treated on equal footing.

Static IR and Raman calculations were performed with a setup analogous to the AIMD-based ones: The frequencies were obtained via normal-mode analysis using the FHI-aims package with the same settings as employed in the AIMD runs. After tightly relaxing the AA' structure, $6N$ structures were generated from the optimized AA' structure ($1\times1\times2$ supercell) by shifting one of the N atoms by 0.001 Å in either $\pm x$, $\pm y$ or $\pm z$ direction, respectively. The forces were computed for each of the $6N$ structures and the elements of the Hessian matrix were evaluated numerically via a central finite difference approach. The Berry phase dipole moments and DFPT polarizabilities for the $6N$ structures were computed with CP2K using again the same settings as for the AIMD snapshots.

AIMD trajectories, input and output files of all calculations, and the raw spectral data are available in the Zenodo database.[102]

4.2 Experimental details

All reagents were acquired from commercial sources and used without further purification. The precursor 1,4-benzene diboronic acid was ordered from Alfa Aesar, while the analytical grade solvents, mesitylene (98%), 1,4-dioxane (≥ 99.8%) and extra dry acetone (99.8%), were purchased from Sigma-Aldrich, Fisher Scientific and Thermo Scientific, respectively.

FT-IR spectra were recorded on a Bruker Vertex 70 FT-IR instrument in the range from 400 cm^{-1} to 4000 cm^{-1} in attenuated total reflectance (ATR) mode. For the measurements, the ATR diamond of the spectrometer was covered with the sample, using approximately 1-2 mg of sample. All measurements were background corrected. Raman measurements were performed on a Bruker RFS 100 FT-Raman spectrometer with a 1064 nm laser and a liquid N_2-cooled Ge-detector. The spectra were background corrected and performed at λ=1064 nm with a resolution of 4 cm^{-1} and 800-1000 mW power (100 scans). For measurement, the sample was packed in a soda lime glass capillary (1.50 mm outer diameter, 1.20 mm inner diameter) inside an Ar-filled glovebox. Details on the PXRD and physisorption measurements are given in Section S3 in the SI, together with the obtained PXRD spectra (Figure S2) and N_2 adsorption-desorption isotherms (Figure S3) for COF-1.

COF-1 was synthesized in accordance with the literature report by Côté and coworkers.[1] A Pyrex tube (1.6 cm outer diameter, 1.4 cm inner diameter) was filled with 1,4-benzene diboronic acid (50 mg, 0.30 mmol), followed by 4 mL solvent mixture of 1,4-dioxane: mesitylene (1:1 v/v) and 2 μL of deionized water. The suspension was then sonicated for 30 min. Next the Pyrex tube was flash frozen at 77 K in liquid N_2 and evacuated to 3 x 10^{-2} mbar before flame-sealing it (length after sealing: 14.8 cm long). The sample was left to react in an oven at 120 °C for 4 days, yielding a white precipitate which was rinsed by filtration through a frit with acetone (30 mL) before activation by CO_2 supercritical drying.

The supercritical drying procedure was performed in a Jumbo Critical Point Dryer 13200JAB (SPI Supplies). Before the supercritical drying process, all samples were soaked in dry acetone for 2 days (solvent was refreshed every 12 h) and placed in POR3 filter crucibles. The drying chamber was first cooled to 17 °C and the filter crucibles containing the samples were placed inside. The chamber was filled with liquid CO_2 and regularly purged for 3 days (multiple times per day). The temperature was then raised to 37 °C, resulting in a chamber pressure of around 80 bar. Once the temperature and pressure were stable, the chamber was left overnight at these conditions. The CO_2 cylinder equipped with a siphon tube was closed, and the pressure was released from the drying chamber. The samples were transferred into an Ar-filled glovebox and handled further under an inert atmosphere.

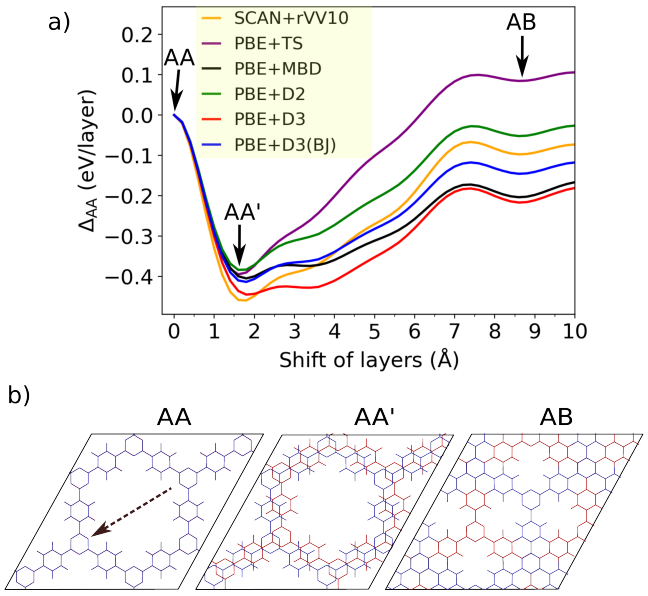

Figure 3 a) Change in the total energy of COF-1 calculated with different London dispersion corrections upon shifting every second layer simultaneously in zigzag direction. The shifted structures corresponding to the AA, AA' (serrated), AB stackings are shown with arrows. Δ_{AA} denotes the change in the energy of the shifted structure with respect to the energy of the AA stacked structure ($\Delta_{AA} = E_{tot}^{AA} - E_{tot}^{Shifted}$). b) The molecular structures for the AA, AA' and AB stacked COF-1. The arrow indicates the direction of the shift.

5 Validation of the computational model

5.1 Dispersion model

The interaction between the layers in materials like COF-1 is dominated by long-range London dispersion interactions. It is well known that generalized gradient approximation (GGA) functionals like PBE cannot describe long-range electron correlation. Empirical dispersion corrections are typically added after the electronic-structure optimization to account for dispersion interactions.[103] We tested five different dispersion models, namely, TS[95] and many-body dispersion (MBD) models[104] as well as Grimme's D2[105] and D3[106] corrections. We investigated also the latter in combination with Becke–Johnson (BJ) damping, denoted as D3(BJ).[107] The dispersion models describe the dispersion at different levels of sophistication. D2 and D3 in-

troduce terms which solely depend on the nuclear coordinates, while TS and MBD also incorporate terms dependent on the electronic density. In addition to these, we also tested the "strongly constrained and appropriately normed" (SCAN) meta-GGA functional [108] combined with the revised Vydrov–van Voorhis (rVV10) non-local correlation functional [109], denoted as SCAN+rVV10 and previously promoted as versatile van der Waals (vdW) functional. [110]

Figure 3a reports the total PES along the zigzag vertices of the hexagonal pores, including electronic contributions at the PBE level and contributions from the dispersion correction, together with the SCAN+rVV10 results. Our starting structure is the density-functional tight binding (DFTB)-optimized AA stacked COF-1 provided by Lukose *et al.* [79] Figure 3a displays the change in the total energy when shifting every second layer with respect to the AA structure as shown in Figure 3b. A shift of 1.6 Å and 8.6 Å corresponds to the AA' and AB stacking, respectively. All models predict that the AA' structure is energetically the most stable one. This aligns with previous DFTB calculations, [28,79] albeit the DFTB-predicted AA' shift is slightly smaller (1.4 – 1.5 Å).

Even though all models agree on AA' being the global minimum, we find that the obtained PES is highly sensitive to the chosen dispersion correction. We observe pronounced qualitative and quantitative differences. For example, PBE+TS estimates for the total energies AB>AA, in agreement with previous DFTB results, [79] while the other five PBE+vdW models and SCAN+rVV10 predict that AA>AB, consistent with a prior PBE+D3(BJ) study. [28] In addition, the quantitative disagreement between SCAN+rVV10, PBE+MBD, PBE+D2, PBE+D3 and PBE+D3(BJ) is large. The calculated energetic differences between AA and AB stacking range from −0.22 eV/layer to 0.08 eV/layer.

More importantly, the dispersion models yield qualitatively different shapes of the PES for shifts close to the global minimum structure AA'. Especially the PBE+D3 model predicts a pronounced second minimum for a shift of 3.6 Å. We found that the double-well-like PES predicted by the PBE+D3 model causes spurious effects in our AIMD simulations. A continuous translation shift was observed throughout the AIMD simulation, where the layers kept moving away from each other. While shear movement of the layers is expected, the stacking between two adjacent layers should be preserved or restored after a couple of picoseconds.

Similarly to the D3 correction, the PBE+MBD and PBE+D3(BJ) models yield an unphysical minimum at 3.6 Å. We also expect in this case that the AIMD trajectory will not remain stable. Conversely, PBE+TS, PBE+D2 and SCAN+rVV10 predict the AA' structure as clear global minimum. The minimum computed with PBE+TS is slightly more pronounced. Hence, we selected the PBE+TS model for our AIMD simulations. As expected, the AA'-type stacking between two adjacent layers was preserved during the AIMD simulations with the PBE+TS model.

Our results show that a proper assessment of the dispersion model against computationally very expensive *ab initio* methods, such as the random phase approximation, [111,112] is in principal necessary. The choice of the dispersion model is expected to have a major impact on the low-frequency modes (< 100 cm^{-1}) because they are dominated by layer-layer interactions. [113] However, we focus in this study on high-frequency modes, which are governed by intralayer interactions. In the latter case, the stability of the AIMD simulation is a sufficient criterion as demonstrated in Sections 6.2 to 6.5.

5.2 Convergence of the supercell size

We assessed the convergence of the electronic PES with respect to supercell and k-mesh sizes. We conducted a similar test as in Section 5.1, shifting the layers as shown in Figure 3b. The electronic PES is reported for different supercell sizes and k-meshes in Figure S6 (SI). We obtain practically identical PESs using a unit cell in combination with a k-point mesh of $1\times1\times3$ (KP setup) and a $1\times1\times2$ supercell (SC setup). However, the KP setup enforces higher symmetry in the z-direction during the AIMD simulation, while the layers 1' and 2' can move independently from the layers 1 and 2 in the SC setup (see Figure 1c for the labels.)

6 IR and Raman results

6.1 Effect of stacking order on the spectra

The KP and SC setups result in a different stacking during the AIMD simulation. While an ordered stacking is enforced in the KP setup, the SC setup allows for a more disordered stacking in z direction, where layers 1 and 2 are no longer aligned on top of 1' and 2', respectively. We assess here the effect of the different stacking on the vibrational bands.

Figure 4b and c compare Raman spectra obtained with the SC and KP setups, respectively, based on polarizabilities calculated with the DFPT approach. Although the same peak positions are obtained with both setups, the intensities are affected. The Raman spectra of the KP setup features sharper and more intense peaks compared to the SC setup, in particular at ≈ 1600 cm^{-1} indicated with the green arrows. However, not all peaks are affected. For example, the intensity of the Raman signal at ≈ 3150 cm^{-1} hardly changes.

To explain the difference between SC and KP intensities, we analyzed the output from our static calculations and assigned the modes. The vibrations occurring at ≈ 3150 cm^{-1} correspond to the C-H stretches, which are isolated and barely affected by the changes in the overall lattice structure. The vibrational modes at ≈ 1600 cm^{-1} are collective motions involving the B-C and C-C stretching modes (Figure

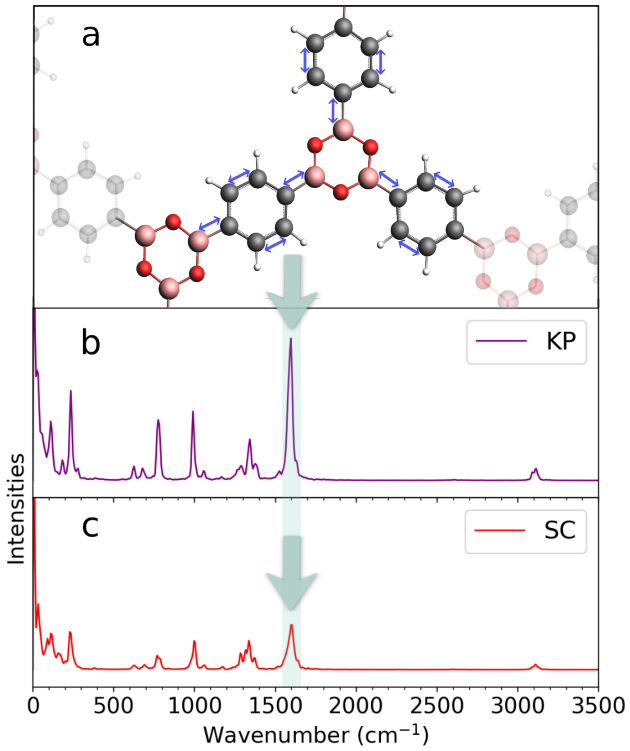

Figure 4 a) COF-1 layer. The arrows indicate the stretching motions of the bonds corresponding to the Raman peaks shown with the green arrows. Raman spectrum calculated from the AIMD trajectory with b) KP setup and c) SC setup.

4a). These vibrations extend across the whole layer. Similarly, the IR signals corresponding to collective vibrational motions appear broader in the SC setup, see Figure S7 (SI). The reason for this observation is that the structure of layers 1 and 1', as well as the 2 and 2', are no longer identical for the snapshots from the SC trajectory. Hence, the collective motions in the layers are slightly different, which manifests in a broadening of the signal.

In the following, we will proceed with the trajectories using the SC setup, motivated by the study already mentioned in Section 3.1, which suggests a statistical stacking model for COF materials.[28] While much larger supercell sizes are necessary to properly describe the disorder in z direction, we consider the AIMD simulations with the SC setup to be the more realistic computational model. Furthermore, the solvent inclusion necessitates an SC setup for accurately modeling the liquid. Thus, we maintain consistency by using the same approach for the pristine material.

6.2 IR spectra

In this section, we investigate the differences between static and MD-based IR spectra, as well as the effects of different dipole approximations on the IR intensities. The IR spectra of the dry COF-1 material from the static and AIMD-based approaches are displayed together with the measured spectrum in Figure 5a. In the static case, the intensities were computed from Berry phase dipole moments. For the MD-based spectra, we tested two different approaches to compute the IR intensities. The total dipole moments were either obtained from the Berry phase approach (MD: μ^{Berry}) or Wannier centers (MD: μ^{MLWF}). The MLWF dipole moments were computed using Equation (12), where the sum runs over all MLWFs in the supercell, with the center of mass of the entire simulation cell taken as the reference point.

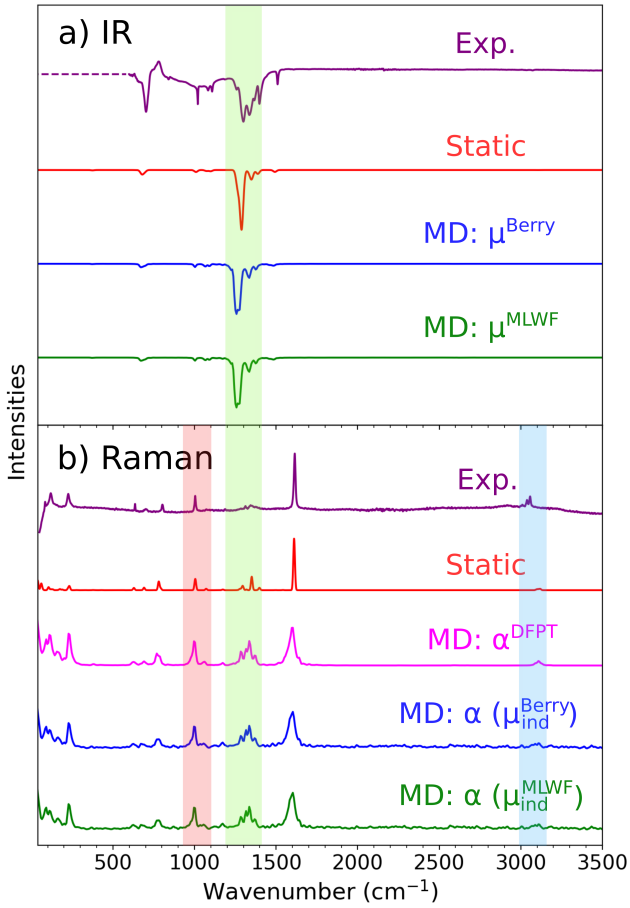

Figure 5 Experimental and computed a) IR and b) Raman spectra for the dry COF-1 material. The computed spectra are either based on the static harmonic approximation or MD. For the MD-based spectra, different approximations to obtain the IR and Raman intensities are displayed. For IR, we computed the intensities either from Berry phase dipole moments (MD: μ^{Berry}) or Wannier centers (MD: μ^{MLWF}). The Raman intensities were computed from DFPT polarizabilities (MD: α^{DFPT}), via the induced dipole method using Berry phase dipoles $\left(\mathrm{MD}: \alpha\left(\mu^{\mathrm{Berry}}_{\mathrm{ind}}\right)\right)$ or Wannier centers $\left(\mathrm{MD}: \alpha\left(\mu^{\mathrm{MLWF}}_{\mathrm{ind}}\right)\right)$.

Figure 5a shows that all calculated spectra reproduce the prominent bands observed in experiment. The bands in the experimental spectrum are generally broader than in the computed spectra, in particular in the range from

$1200 - 1400$ cm^{-1} (area highlighted in green), where most of the collective vibrations involving B-C and B-O stretching modes occur. The most notable difference is the band at $\approx$ 650 cm^{-1}. While this band is also present in the computed spectra, the intensity is much higher in the experimental one.

Comparing the computed static and Berry phase MD-based spectra, we find that both agree well in terms of band intensities and frequencies. The MD-based spectrum has a broader band at $1200 - 1400$ cm^{-1}, which must be attributed to the stacking disorder which is incorporated to some extent in our MD calculations with SC setup; see Section 6.1.

Next, we compare the Berry phase and MLWF-based IR spectra. The Berry phase and the MLWF-based approach yield exactly the same spectra. This is expected, and it can be ascribed to the fact that under PBCs, Wannier centers are the real-space equivalent of the k-space Berry phase expression.[68]

6.3 Raman spectra

Similarly as in Section 6.2, we compare the computed static and MD-based Raman spectra to experiment. The Raman spectra for the dry COF-1 are presented in Figure 5b. The intensities of the static Raman spectrum were computed from DFPT polarizabilities. In the MD case, we employ three different methods. These methods include the polarizability tensors calculated from induced dipole moments (Equation (13)), denoted as $\alpha(\mu_{\text{ind}})$, and the DFPT approach (Equation (17)), denoted as α^{DFPT}. We computed the induced dipole moments using the Berry phase scheme and MLWFs.

In general, all computed spectra agree well with the experimental Raman spectrum. However, the static approach fails to reproduce the signals below 250 cm^{-1}. These features are reproduced in all MD-based Raman spectrum. Similarly to the IR spectra, the MD-based spectra generally exhibit broader peaks.

Turning now to the different approximations used for the polarizability tensor α, we find that α^{DFPT} and $\alpha\left(\mu_{\text{ind}}^{\text{Berry}}\right)$ yield almost identical MD-based spectra. The good agreement is not surprising since $\alpha\left(\mu_{\text{ind}}^{\text{Berry}}\right)$ can be viewed as the non-analytic variant of α^{DFPT}. The derivative $\partial\mu_\alpha/\partial E_\beta$ (Equation (17)) is computed analytically in DFPT and numerically in the induced dipole approach. In both cases, μ_α corresponds to a Berry phase dipole moment. The comparison of the $\alpha\left(\mu_{\text{ind}}^{\text{Berry}}\right)$ and $\alpha\left(\mu_{\text{ind}}^{\text{MLWF}}\right)$ reveals that, in line with the observations for IR, we obtain identical spectra from the Berry phase and MLWF-based approaches.

In general, we find that the induced dipole spectra exhibit more noise, making smaller peaks such as the C-H stretches around $\approx$ 3150 cm^{-1} (area highlighted in blue) hardly visible. Our tests revealed that the induced dipole moments are highly sensitive to the strength of the applied electric field. Even a decrease by $\approx$ 0.0045 a.u. in the electric field intensity E can introduce so much noise that no peaks are visible at all. Using forward instead of central finite differences might be another, but probably less important, factor contributing to the noise in the spectra. We note that a central-finite-difference approach would double the number of induced-field calculations.

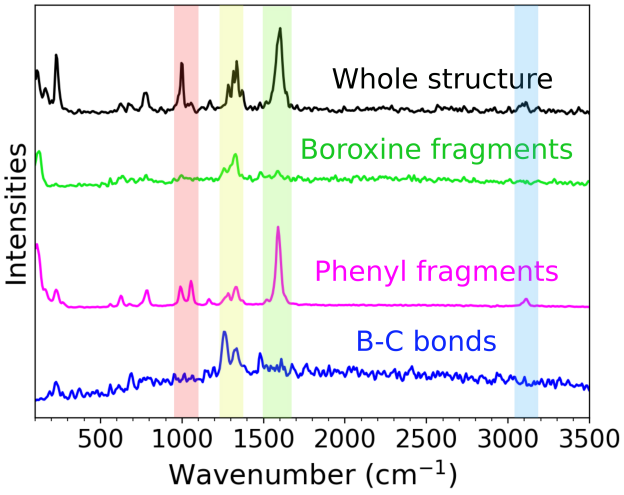

Figure 6 The AIMD-based Raman spectra for pristine COF-1 calculated from MLWFs. The subspectra of the boroxine/phenyl units and B-C bonds are calculated by defining fragments as explained in Section 3.2.1.

The MD-based α^{DFPT} approach yields the best results compared to experiment and will be used as default unless otherwise stated. The DFPT implementation in CP2K is computationally very efficient, with the computational effort not significantly exceeding the time required to perform three additional sets of electric field calculations for each snapshot, as necessary in the induced-field method.

6.4 Dissecting the COF-1 spectra

In this section, we demonstrate how the MLWF-based fragment approach introduced in Section 3.2.1 can assist in dissecting the AIMD-based spectra into individual contributions of the phenyl and boroxine rings and B-C bonds. The dissected Raman spectrum generated from $\alpha\left(\mu_{\text{ind}}^{\text{MLWF}}\right)$ is shown in Figure 6 and the IR spectrum in Figure S8 (SI).

Figure 6 indicates that most peaks in the spectrum of the COF-1 structure can easily be assigned to the phenyl and boroxine contributions. Starting with the area highlighted in blue, we know from normal mode analysis that the peak at 3150 cm^{-1} originates from the C-H stretches. We can now also confirm with the AIMD-based approach that this band is indeed associated with the phenyl fragments. Similarly, the peak at 1600 cm^{-1} (area highlighted in green) clearly originates from the phenyl fragments. Normal mode analysis also shows that this band belongs to the aromatic C-C stretches.

The B-C bond stretches appear in the range of 1250-

1400 cm^{-1} (area highlighted in yellow). This observation is also in line with the normal mode analysis results predicting B-C stretches exactly in this region. We observe that boroxine and phenyl fragments also show peaks in this area possibly because the B-C bonds stretches are collective vibrational modes which affect the overall structure.

The peak at $\approx$ 1000 cm^{-1} (area highlighted in red) appears to result from vibrations of the phenyl ring. Normal mode analysis suggests that this peak corresponds to the aromatic C-C bending modes. Notably, the spectrum of the phenyl fragments features two peaks in the region highlighted in red, rather than the single prominent peak visible in the spectrum of the whole structure. Artificial peaks have been previously observed in AIMD-computed vibrational spectra of liquid benzene using MLWFs.[82] This observation was explained as follows and might also apply here: Throughout the AIMD simulation, the locations of the MLWFs representing the double and single bonds of the phenyl ring interchange. This leads to spurious changes of the dipole moment, and eventually artificial bands in the spectrum. Artificial peaks can also be observed in the low-frequency region of the MD-based IR spectrum of the phenyl fragments (Figure S8, SI). However, the effect is small and does not hamper the assignment of the peaks in the overall vibrational spectra to boroxine or phenyl fragments.

The identification of vibrational bands associated with the covalent bonds connecting the molecular fragments is often particularly valuable when the framework material is synthesized via a bimolecular reaction mechanism. Such an example could be the metal-phthalocyanine-based pyrazine-linked conjugated 2D COFs that are sythesized via a condensation reaction to form pyrazine linkages which are not present in any of the reactants.[114] However, COF-1 is synthesized via a unimolecular reaction, where the B-C bond is already present in the precursor molecule (1,4-benzene diboronic acid), making the monitoring of this bond in *in situ* Raman spectroscopy not beneficial. The purpose here is simply to demonstrate that such an analysis is possible.

6.5 Solvent signature

So far, we presented only vibrational spectra of the pristine COF-1 material. We made a dedicated effort to dry the sample to enable a comparison between theoretical predictions and experimental results. If such an effort is not made, the samples are not completely dry and include residual solvent, in our case 1,4-dioxane and mesitylene, trapped inside the pores.[1] In the following, we discuss the effect of the solvent residues, or guest molecules, on IR and Raman spectra, using the computational approach and settings we determined as optimal in Sections 6.2 and 6.3 (Berry phase dipole moments and DFPT polarizabilities).

The experimental and MD-based IR and Raman spectra for the dry (without solvent) and wet (with solvent) COF-1 system are presented in Figures 7a and b, respectively. In addition, we computed the contribution of the 1,4-dioxane solvent molecules to the overall MD-based spectra. The solvent contribution was extracted by computing the molecular dipoles of the 1,4-dioxane molecules using MLWFs for all selected snapshots of the COF-1+solvent trajectory. Subsequently, the IR and Raman spectrum were generated solely from the molecular dipole moments of the solvent.

Starting with the measured IR spectra displayed in Figure 7a, we observe additional bands for the wet sample at following three wavenumbers 800 – 900 cm^{-1}, 1600 cm^{-1}, and, arguably, at 2850 – 3050 cm^{-1}, which are not present in the spectrum of the dry COF-1 sample. These bands are highlighted in yellow, green, and blue, respectively, in Figure 7a. Otherwise, both measured spectra are very similar.

Comparing the computed MD-based IR spectra with and without solvent, the additional bands at 800 – 900 cm^{-1} occur also in the simulated spectrum of COF-1+solvent, albeit with much lower intensity than in experiment. Normal mode analysis of isolated molecules reveals that these bands belong to the C-O bending modes of 1,4-dioxane. The additional bands at 2850 – 3050 cm^{-1} also appear in the computed spectrum, but, similarly as in experiment, with very low intensity. These bands correspond to the C-H stretching modes of 1,4-dioxane. The third small band at 1600 cm^{-1} is not present in our simulation. Vibrations of COF-1 or 1,4-dioxane are neither computationally or experimentally expected at this wavenumber. This band most likely originates either from the precursor 1,4-benzene diboronic acid (BDBA) (see Figure S9a in the SI) or the second solvent used throughout the synthesis, namely mesitylene.

The most prominent difference we observe between the two MD-based IR spectra is that the broad band at 1250 – 1400 cm^{-1} exhibits higher intensity in the COF-1+solvent spectrum compared to the pristine COF-1 spectrum. One possible explanation for the enhancement of the intensity of this band could be that the presence of the solvent molecules inside the COF-1 pores restricts the movement of the layers. This preservation of stacking symmetry during the MD simulation may result in narrower peaks.

Turning now to Raman and comparing the measured spectrum of the dry and wet sample, we observe additional peaks at $\approx$ 500-600 cm^{-1} and $\approx$ 2850-3050 cm^{-1}, highlighted in red and blue in Figure 7b, respectively, for the wet sample. Apart from these, the Raman spectrum remains unchanged. The AIMD-based spectrum of COF-1+solvent does not show additional signals at $\approx$ 500-600 cm^{-1} in Figure 7b, indicating that these peaks might stem from the remaining mesitylene or the precursor. However, the peaks at $\approx$ 2850-3050 are reproduced. In agreement with the experimental findings, the solvent has no further impact on the spectrum, except that the peaks involving collective vibrations, such as the one at 1600 cm^{-1}, appear narrower in the

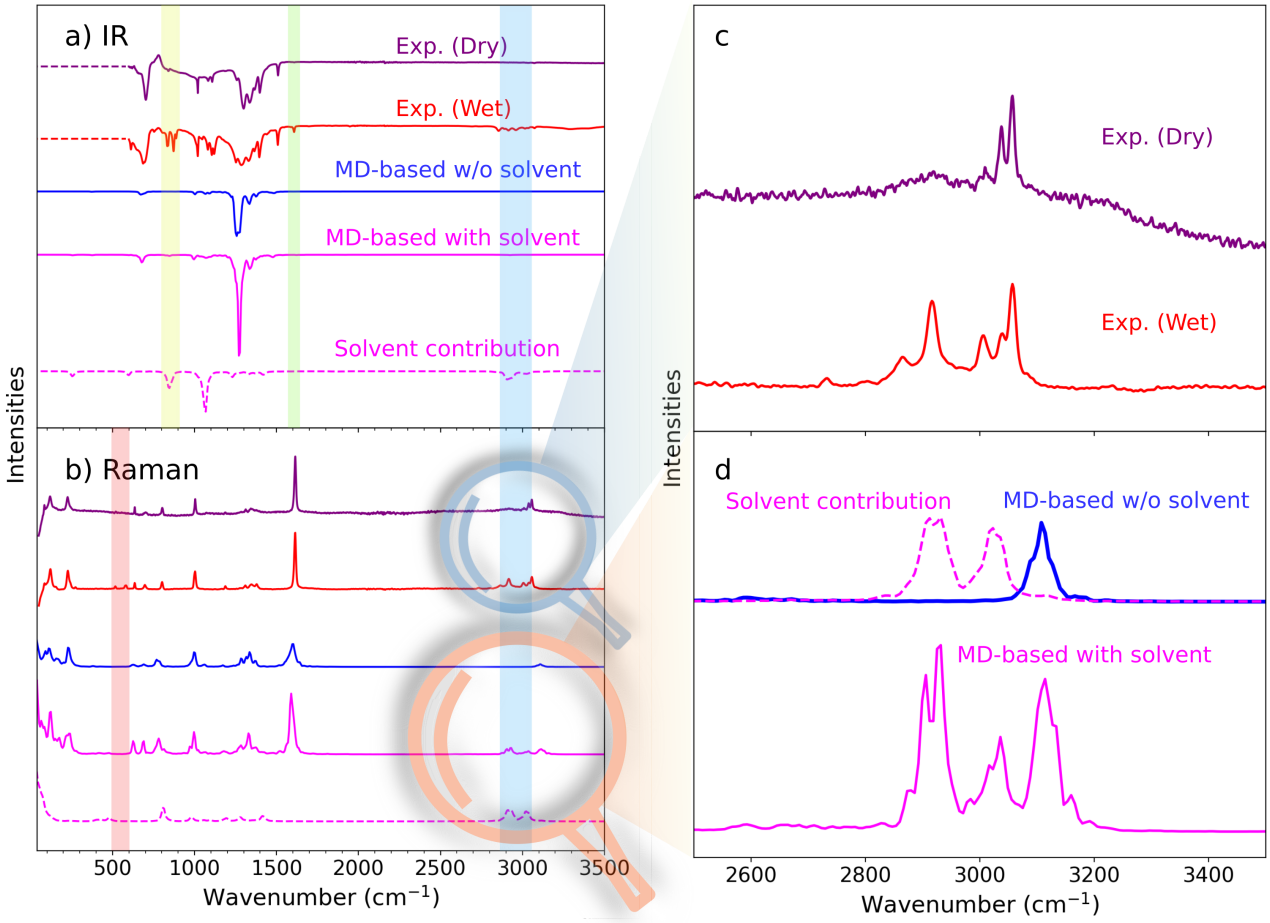

Figure 7 a) IR and b) Raman spectra of COF-1 comparing experimental data and computed spectra from the MD-based approach. The MD-based spectra were computed for the dry COF-1 (without solvent, blue line) and with solvent (1,4-dioxane, pink line). The "wet" experimental spectra is expected to contain also vibrational excitations of the 1,4-dioxane solvent. The spectra of the solvent contribution were obtained from the COF-1+solvent trajectory by extracting solvent contributions via MLWFs. (c) Zoom into frequency range of 2500-3500 cm^{-1} for experimental Raman spectra with visible solvent contributions at around 2850 cm^{-1} and 3050 cm^{-1}. (d) Zoom into the same frequency region for the MD-based spectra, displaying the dissection of the MD-based Raman spectrum in COF-1 and solvent contribution.

spectra containing solvent. This observation aligns with the findings for IR. The explanation is the same.

We take the vibrational signature of the solvent in the range from 2850-3050 cm^{-1} as an example to illustrate how the MD-based approach can aid the peak assignment in experimental spectra. The zoom-ins are shown in Figure 7c and d for measured and simulated Raman spectra, respectively. Comparing the measured Raman spectra of dry and wet sample, two additional spectral features are observed at $\approx$ 2900 and 3000 cm^{-1} for the wet sample. These peaks could be originating from the remaining precursor BDBA (see also Figure S9b in SI) or the solvent molecules trapped inside the pores. The simulated spectra now clearly prove that these two additional spectral features must originate from the 1,4-dioxane solvent. Normal mode analysis of isolated 1,4-dioxane molecules suggest that those must be attributed to C-H symmetric and asymmetric stretching modes. The C-H stretching modes of COF-1 appear closer to 3200 cm^{-1}.

7 Conclusions

In this study, we presented the IR and Raman spectra for COF-1, a well-known 2D periodic layered material, using an AIMD-based procedure. We compared our results to static calculations and experimental data. Our computational studies comprised the pristine COF-1 material without solvent, as well as COF-1 with 1,4-dioxane solvent molecules included in the pores.

We carefully validated our computational model regarding the dispersion correction, supercell and k-mesh sizes, and stacking order. After investigating several dispersion models, we conclude that the choice of dispersion model must be made with great care, which very likely is also true for other layered materials. We observed that the TS correction,

unlike most other models, yields a physically sound PES for COF-1, resulting in well-preserved layer stackings throughout the AIMD simulation. Furthermore, our findings indicate that the stacking order influences the width and intensity of certain peaks. We recommend using a supercell rather than a *k*-mesh setup to account for stacking disorder between non-adjacent layers, as found in a recent previous study.[28]

We assessed different methods to compute dipole moments and polarizability tensors and their effect on AIMD-based IR and Raman intensities. We found that Berry phase dipole moments and DFPT polarizabilities yield the best agreement with experiment for the overall IR and Raman spectra, respectively. We also demonstrated that MLWFs can be used to analyse and dissect the AIMD-computed spectra into contributions from individual molecular building blocks and chemical bonds. This approach can possibly be used in *in situ* monitoring of reactions of other porous materials which form via bimolecular mechanisms, since it allows tracking the formation of the bonds connecting the reactants.

Finally, we investigated the host-guest interactions between COF-1 and solvent molecules in measured and AIMD-computed spectra. The influence of the solvent on COF-1 spectra is relatively minor, with the most significant impact observed in Raman spectra at high wavenumbers (≈ 3000 cm^{-1}). The AIMD-based spectra correctly reproduce these observations and confirm that the additional peaks originate from 1,4-dioxane.

Future and ongoing work comprises a dedicated investigation of low-frequency modes in 2D organic polymers and inorganic 2D materials as well as the acceleration of the MD-based procedure using machine-learned potentials. Another focus is the extension to resonance Raman spectroscopy.

Acknowledgement All authors thank the German Research Foundation for support within the CRC 1415 and RTG 2861. The ZIH of the TU Dresden, the Jülich Supercomputing Centre, the Paderborn Center for Parallel Computing (PC2) as well as the Finnish CSC - IT Center for Science are gratefully acknowledged for providing computational resources. A.S. and K.M.G.A. gratefully acknowledge the Fonds der Chemischen Industrie for a Liebig Fellowship and a Doktorandenstipendium. D.G. acknowledges the Emmy Noether Programme of the German Research Foundation (project number 453275048) for funding. We also thank Martin Brehm for helpful discussions.

Supporting Information Available

The supplementary information is available free of charge. We give the key equations for the static IR and Raman intensities which rely on the normal mode analysis within the harmonic approximation (Section S1). We include the details on how we calculated the power spectra (Section S2), together with the power spectrum of COF-1 (Figure S1). We provide the details on the PXRD and physisorption measurements in Section S3, including the PXRD spectra of the dry and wet COF-1 samples (Figure S2) and the N_2 adsorption-desorption isotherms of the dry COF-1 sample (Figure S3). We present the power, IR and Raman spectra of liquid water calculated with Travis program package and our own methodology (Figure S4). We show the PES diagrams of COF-1 calculated using different codes (CP2K and FHI-aims), setups (KP and SC), supercell sizes and *k*-point meshes (Figures S5 and S6). We show the IR spectra of the SC and KP setups of COF-1 calculated from the Berry phase dipole moments (Figure S7). We include the AIMD-based IR spectra showing contributions from the boroxine and phenyl fragments and B-C bonds in Figure S8. We also present the calculated IR (Figure S9a) and the experimental and calculated Raman spectra (Figure S9b) of the precursor BDBA.

References

(1) Côté, A. P.; Benin, A. I.; Ockwig, N. W.; O'Keeffe, M.; Matzger, A. J.; Yaghi, O. M. Porous, crystalline, covalent organic frameworks. *Science* **2005**, *310*, 1166–1170.

(2) Feng, X.; Ding, X.; Jiang, D. Covalent organic frameworks. *Chem. Soc. Rev.* **2012**, *41*, 6010–6022.

(3) Han, S. S.; Furukawa, H.; Yaghi, O. M.; Goddard, W. A. I. Covalent Organic Frameworks as Exceptional Hydrogen Storage Materials. *J. Am. Chem. Soc.* **2008**, *130*, 11580–11581.

(4) Furukawa, H.; Yaghi, O. M. Storage of hydrogen, methane, and carbon dioxide in highly porous covalent organic frameworks for clean energy applications. *J. Am. Chem. Soc.* **2009**, *131*, 8875–8883.

(5) Fan, H.; Mundstock, A.; Feldhoff, A.; Knebel, A.; Gu, J.; Meng, H.; Caro, J. Covalent organic framework–covalent organic framework bilayer membranes for highly selective gas separation. *J. Am. Chem. Soc.* **2018**, *140*, 10094–10098.

(6) Ying, Y.; Tong, M.; Ning, S.; Ravi, S. K.; Peh, S. B.; Tan, S. C.; Pennycook, S. J.; Zhao, D. Ultrathin two-dimensional membranes assembled by ionic covalent organic nanosheets with reduced apertures for gas separation. *J. Am. Chem. Soc.* **2020**, *142*, 4472–4480.

(7) Wang, Z.; Zhang, Y.; Lin, E.; Geng, S.; Wang, M.; Liu, J.; Chen, Y.; Cheng, P.; Zhang, Z. Kilogram-scale

fabrication of a robust olefin-linked covalent organic framework for separating ethylene from a ternary C2 hydrocarbon mixture. *J. Am. Chem. Soc.* **2023**, *145*, 21483–21490.

(8) De, A.; Haldar, S.; Schmidt, J.; Amirjalayer, S.; Reichmayr, F.; Lopatik, N.; Shupletsov, L.; Brunner, E.; Weidinger, I. M.; Schneemann, A. An Alkyne-Bridged Covalent Organic Framework Featuring Interactive Pockets for Bromine Capture. *Angew. Chem. Int. Ed.* **2024**, e202403658.

(9) Dalapati, S.; Jin, S.; Gao, J.; Xu, Y.; Nagai, A.; Jiang, D. An azine-linked covalent organic framework. *J. Am. Chem. Soc.* **2013**, *135*, 17310–17313.

(10) Das, G.; Biswal, B. P.; Kandambeth, S.; Venkatesh, V.; Kaur, G.; Addicoat, M.; Heine, T.; Verma, S.; Banerjee, R. Chemical sensing in two dimensional porous covalent organic nanosheets. *Chem. Sci.* **2015**, *6*, 3931–3939.

(11) Chandra, S.; Kundu, T.; Kandambeth, S.; BabaRao, R.; Marathe, Y.; Kunjir, S. M.; Banerjee, R. Phosphoric Acid Loaded Azo (- N=N-) Based Covalent Organic Framework for Proton Conduction. *J. Am. Chem. Soc.* **2014**, *136*, 6570–6573.

(12) Chandra, S.; Kundu, T.; Dey, K.; Addicoat, M.; Heine, T.; Banerjee, R. Interplaying intrinsic and extrinsic proton conductivities in covalent organic frameworks. *Chem. Mater.* **2016**, *28*, 1489–1494.

(13) Ding, S.-Y.; Gao, J.; Wang, Q.; Zhang, Y.; Song, W.-G.; Su, C.-Y.; Wang, W. Construction of covalent organic framework for catalysis: Pd/COF-LZU1 in Suzuki–Miyaura coupling reaction. *J. Am. Chem. Soc.* **2011**, *133*, 19816–19822.

(14) Xu, H.; Chen, X.; Gao, J.; Lin, J.; Addicoat, M.; Irle, S.; Jiang, D. Catalytic covalent organic frameworks via pore surface engineering. *Chem. Comm.* **2014**, *50*, 1292–1294.

(15) Xu, F.; Jin, S.; Zhong, H.; Wu, D.; Yang, X.; Chen, X.; Wei, H.; Fu, R.; Jiang, D. Electrochemically active, crystalline, mesoporous covalent organic frameworks on carbon nanotubes for synergistic lithium-ion battery energy storage. *Sci. Rep.* **2015**, *5*, 8225.

(16) Haldar, S. et al. Porous dithiine-linked covalent organic framework as a dynamic platform for covalent polysulfide anchoring in lithium–sulfur battery cathodes. *J. Am. Chem. Soc.* **2022**, *144*, 9101–9112.

(17) Liu, Y. et al. Redox-Bipolar Polyimide Two-Dimensional Covalent Organic Framework Cathodes for Durable Aluminium Batteries. *Angew. Chem. Int. Ed.* **2023**, *62*, e202306091.

(18) Haldar, S.; Waentig, A. L.; Ramuglia, A. R.; Bhauriyal, P.; Khan, A. H.; Pastoetter, D. L.; Isaacs, M. A.; De, A.; Brunner, E.; Wang, M.; others Covalent Trapping of Cyclic-Polysulfides in Perfluorinated Vinylene-Linked Frameworks for Designing Lithium-Organosulfide Batteries. *ACS Energy Lett.* **2023**, *8*, 5098–5106.

(19) Ding, X.; Guo, J.; Feng, X.; Honsho, Y.; Guo, J.; Seki, S.; Maitarad, P.; Saeki, A.; Nagase, S.; Jiang, D. Synthesis of metallophthalocyanine covalent organic frameworks that exhibit high carrier mobility and photoconductivity. *Angew. Chem. Int. Ed.* **2011**, *6*, 1289–1293.

(20) Jin, E.; Lan, Z.; Jiang, Q.; Geng, K.; Li, G.; Wang, X.; Jiang, D. 2D sp2 carbon-conjugated covalent organic frameworks for photocatalytic hydrogen production from water. *Chem* **2019**, *5*, 1632–1647.

(21) Wan, S.; Guo, J.; Kim, J.; Ihee, H.; Jiang, D. A belt-shaped, blue luminescent, and semiconducting covalent organic framework. *Angew. Chem. Int. Ed.* **2008**, *47*, 8826–8830.

(22) Jin, S.; Ding, X.; Feng, X.; Supur, M.; Furukawa, K.; Takahashi, S.; Addicoat, M.; El-Khouly, M. E.; Nakamura, T.; Irle, S.; Fukuzumi, S.; Nagai, A.; Jiang, D. Charge dynamics in a donor-acceptor covalent organic framework with periodically ordered bicontinuous heterojunctions. *Angew. Chem. Int. Ed.* **2013**, *52*, 2017–2021.

(23) Wang, M.; Fu, S.; Petkov, P.; Fu, Y.; Zhang, Z.; Liu, Y.; Ma, J.; Chen, G.; Gali, S. M.; Gao, L.; others Exceptionally high charge mobility in phthalocyanine-based poly (benzimidazobenzophenanthroline)-ladder-type two-dimensional conjugated polymers. *Nat. Mater.* **2023**, *22*, 880–887.

(24) Guo, Y.; Guo, S.; Ren, J.; Zhai, Y.; Dong, S.; Wang, E. Cyclodextrin functionalized graphene nanosheets with high supramolecular recognition capability: synthesis and host- guest inclusion for enhanced electrochemical performance. *ACS nano* **2010**, *4*, 4001–4010.

(25) Sahabudeen, H.; Qi, H.; Glatz, B. A.; Tranca, D.; Dong, R.; Hou, Y.; Zhang, T.; Kuttner, C.; Lehnert, T.; Seifert, G.; Kaiser, U.; Fery, A.; Zheng, Z.; Feng, X.

Wafer-sized multifunctional polyimine-based two-dimensional conjugated polymers with high mechanical stiffness. *Nat. Commun.* **2016**, *7*, 13461.

(26) Liu, K.; Qi, H.; Dong, R.; Shivhare, R.; Addicoat, M.; Zhang, T.; Sahabudeen, H.; Heine, T.; Mannsfeld, S.; Kaiser, U.; Zheng, Z.; Feng, X. On-water surface synthesis of crystalline, few-layer two-dimensional polymers assisted by surfactant monolayers. *Nat. Chem.* **2019**, *11*, 994–1000.

(27) Du, Y.; Calabro, D.; Wooler, B.; Li, Q.; Cundy, S.; Kamakoti, P.; Colmyer, D.; Mao, K.; Ravikovitch, P. Kinetic and mechanistic study of COF-1 phase change from a staggered to eclipsed model upon partial removal of mesitylene. *J. Phys. Chem. C* **2014**, *118*, 399–407.

(28) Zhang, Y.; Polozij, M.; Heine, T. Statistical representation of stacking disorder in layered covalent organic frameworks. *Chem. Mater.* **2022**, *34*, 2376–2381.

(29) Sun, J.; Iakunkov, A.; Baburin, I. A.; Joseph, B.; Palermo, V.; Talyzin, A. V. Covalent Organic Framework (COF-1) under High Pressure. *Angew. Chem. Int. Ed.* **2020**, *59*, 1087–1092.

(30) Zhong, H.; Wang, M.; Ghorbani-Asl, M.; Zhang, J.; Ly, K. H.; Liao, Z.; Chen, G.; Wei, Y.; Biswal, B. P.; Zschech, E.; Weidinger, I. M.; Krasheninnikov, A. V.; Dong, R.; Feng, X. Boosting the Electrocatalytic Conversion of Nitrogen to Ammonia on Metal-Phthalocyanine-Based Two-Dimensional Conjugated Covalent Organic Frameworks. *J. Am. Chem. Soc.* **2021**, *143*, 19992–20000, PMID: 34784212.

(31) Ly, K. H.; Weidinger, I. M. Understanding active sites in molecular (photo) electrocatalysis through complementary vibrational spectroelectrochemistry. *Chem. Commun.* **2021**, *57*, 2328–2342.

(32) Long, D. A. *The Raman Effect*; John Wiley & Sons Ltd, 2002.

(33) Redington, R. L.; Redington, T. E.; Sams, R. L. Quantum tunneling in the midrange vibrational fundamentals of tropolone. *J. Phys. Chem. A* **2006**, *110*, 9633–9642.

(34) Giese, K.; Petković, M.; Naundorf, H.; Kühn, O. Multidimensional quantum dynamics and infrared spectroscopy of hydrogen bonds. *Phys. Rep.* **2006**, *430*, 211–276.

(35) Silvestrelli, P. L.; Bernasconi, M.; Parrinello, M. Ab initio infrared spectrum of liquid water. *Chem. Phys. Lett.* **1997**, *277*, 478–482.

(36) Bernasconi, M.; Silvestrelli, P.; Parrinello, M. Ab initio infrared absorption study of the hydrogen-bond symmetrization in ice. *Phys. Rev. Lett.* **1998**, *81*, 1235.

(37) Praprotnik, M.; Janežič, D.; Mavri, J. Temperature Dependence of Water Vibrational Spectrum: A Molecular Dynamics Simulation Study. *J. Phys. Chem. A* **2004**, *108*, 11056–11062.

(38) Iftimie, R.; Tuckerman, M. E. Decomposing total IR spectra of aqueous systems into solute and solvent contributions: A computational approach using maximally localized Wannier orbitals. *J. Chem. Phys.* **2005**, *122*, 214508.

(39) Thomas, M.; Brehm, M.; Fligg, R.; Vöhringer, P.; Kirchner, B. Computing vibrational spectra from ab initio molecular dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608–6622.

(40) Luber, S.; Iannuzzi, M.; Hutter, J. Raman spectra from ab initio molecular dynamics and its application to liquid *S*-methyloxirane. *J. Chem. Phys.* **2014**, *141*, 094503.

(41) Taherivardanjani, S.; Elfgen, R.; Reckien, W.; Suarez, E.; Perlt, E.; Kirchner, B. Benchmarking the Computational Costs and Quality of Vibrational Spectra from Ab Initio Simulations. *Adv. Theory Simul.* **2022**, *5*, 2100293.

(42) Brehm, M.; Thomas, M. Computing bulk phase resonance Raman spectra from ab initio molecular dynamics and real-time TDDFT. *J. Chem. Theory Comput.* **2019**, *15*, 3901–3905.

(43) Wathelet, V.; Champagne, B.; Mosley, D. H.; André, J.-M.; Massidda, S. Vibrational frequencies of diatomic molecules from Car and Parrinello molecular dynamics. *Chem. Phys. Lett.* **1997**, *275*, 506–512.

(44) Shang, H.; Raimbault, N.; Rinke, P.; Scheffler, M.; Rossi, M.; Carbogno, C. All-electron, real-space perturbation theory for homogeneous electric fields: theory, implementation, and application within DFT. *New J. Phys.* **2018**, *20*, 073040.

(45) Mattiat, J.; Luber, S. Efficient calculation of (resonance) Raman spectra and excitation profiles with real-time propagation. *J. Chem. Phys.* **2018**, *149*, 174108.

(46) Litman, Y.; Behler, J.; Rossi, M. Temperature dependence of the vibrational spectrum of porphycene: a qualitative failure of classical-nuclei molecular dynamics. *Farad. Disc.* **2020**, *221*, 526–546.

(47) Morita, A.; Hynes, J. T. A theoretical analysis of the sum frequency generation spectrum of the water surface. *Chem. Phys.* **2000**, *258*, 371–390.

(48) Litman, Y.; Lan, J.; Nagata, Y.; Wilkins, D. M. Fully first-principles surface spectroscopy with machine learning. *J. Phys. Chem. Lett.* **2023**, *14*, 8175–8182.

(49) Litman, Y.; Chiang, K.-Y.; Seki, T.; Nagata, Y.; Bonn, M. Surface stratification determines the interfacial water structure of simple electrolyte solutions. *Nat. Chem.* **2024**, 1–7.

(50) Putrino, A.; Parrinello, M. Anharmonic Raman spectra in high-pressure ice from ab initio simulations. *Phys. Rev. Lett.* **2002**, *88*, 176401.

(51) Buch, V.; Mohamed, F.; Parrinello, M.; Devlin, J. Elusive structure of HCl monohydrate. *J. Chem. Phys.* **2007**, *126*, 074503.

(52) Buch, V.; Mohamed, F.; Parrinello, M.; Devlin, J. A new glance at HCl-monohydrate spectroscopy, using on-the-fly dynamics. *J. Chem. Phys.* **2007**, *126*, 021102.

(53) Pagliai, M.; Cavazzoni, C.; Cardini, G.; Erbacci, G.; Parrinello, M.; Schettino, V. Anharmonic infrared and Raman spectra in Car–Parrinello molecular dynamics simulations. *J. Chem. Phys.* **2008**, *128*, 224514.

(54) Raimbault, N.; Grisafi, A.; Ceriotti, M.; Rossi, M. Using Gaussian process regression to simulate the vibrational Raman spectra of molecular crystals. *New J. Phys.* **2019**, *21*, 105001.

(55) Beynon, O. T.; Owens, A.; Carbogno, C.; Logsdail, A. J. Evaluating the role of anharmonic vibrations in zeolite β materials. *J. Phys. Chem. C* **2023**, *127*, 16030–16040.

(56) Hoffman, A. E.; Temmerman, W.; Campbell, E.; Damin, A. A.; Lezcano-Gonzalez, I.; Beale, A. M.; Bordiga, S.; Hofkens, J.; Van Speybroeck, V. A Critical Assessment on Calculating Vibrational Spectra in Nanostructured Materials. *J. Chem. Theory Comput.* **2023**, *20*, 513–531.

(57) Zhu, L.; Zeng, X.; Shang, H.; Li, Z. First-Principles Analysis of the Raman Spectra of 2D Material YbOCl. *J. Phys. Chem. C* **2023**, *127*, 23359–23369.

(58) Berger, E.; Lv, Z.-P.; Komsa, H.-P. Raman spectra of 2D titanium carbide MXene from machine-learning force field molecular dynamics. *J. Mater. Chem. C* **2023**, *11*, 1311–1319.

(59) Thomas, M. *Theoretical modeling of vibrational spectra in the liquid phase*; Springer, 2016.

(60) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 2017.

(61) Gordon, R. Molecular motion in infrared and Raman spectra. *J. Chem. Phys.* **1965**, *43*, 1307–1312.

(62) McQuarrie, D. A. *Quantum Chemistry*; University Science Books, 2008.

(63) King-Smith, R. D.; Vanderbilt, D. Theory of polarization of crystalline solids. *Phys. Rev. B* **1993**, *47*, 1651–1654.

(64) Resta, R. Macroscopic Electric Polarization as a Geometric Quantum Phase. *Europhys. Lett.* **1993**, *22*, 133.

(65) Vanderbilt, D.; King-Smith, R. D. Electric polarization as a bulk quantity and its relation to surface charge. *Phys. Rev. B* **1993**, *48*, 4442–4455.

(66) Vanderbilt, D. *Berry phases in electronic structure theory: electric polarization, orbital magnetization and topological insulators*; Cambridge University Press, 2018.

(67) Marzari, N.; Vanderbilt, D. Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B* **1997**, *56*, 12847.

(68) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. Maximally localized Wannier functions: Theory and applications. *Rev. Mod. Phys.* **2012**, *84*, 1419–1475.

(69) Berghold, G.; Mundy, C. J.; Romero, A. H.; Hutter, J.; Parrinello, M. General and efficient algorithms for obtaining maximally localized Wannier functions. *Phys. Rev. B* **2000**, *61*, 10040–10048.

(70) Schreder, L.; Luber, S. Local approaches for electric dipole moments in periodic systems and their application to real-time time-dependent density functional theory. *J. Chem. Phys.* **2021**, *155*, 134116.

(71) Souza, I.; Íniguez, J.; Vanderbilt, D. First-principles approach to insulators in finite electric fields. *Phys. Rev. Lett.* **2002**, *89*, 117602.

(72) Umari, P.; Pasquarello, A. Ab initio molecular dynamics in a finite homogeneous electric field. *Phys. Rev. Lett.* **2002**, *89*, 157602.

(73) Gonze, X.; Vigneron, J.-P. Density-functional approach to nonlinear-response coefficients of solids. *Phys. Rev. B* **1989**, *39*, 13120.

(74) Gonze, X. Adiabatic density-functional perturbation theory. *Phys. Rev. A* **1995**, *52*, 1096.

(75) Gerratt, J.; Mills, I. M. Force Constants and Dipole-Moment Derivatives of Molecules from Perturbed Hartree–Fock Calculations. I. *J. Chem. Phys.* **1968**, *49*, 1719–1729.

(76) Ochsenfeld, C.; Head-Gordon, M. A reformulation of the coupled perturbed self-consistent field equations entirely within a local atomic orbital density matrix-based scheme. *Chem. Phys. Lett.* **1997**, *270*, 399–405.

(77) Marx, D.; Hutter, J. *Ab initio molecular dynamics: basic theory and advanced methods*; Cambridge University Press, 2009.

(78) Sun, J.; Klechikov, A.; Moise, C.; Prodana, M.; Enachescu, M.; Talyzin, A. V. A molecular pillar approach to grow vertical covalent organic framework nanosheets on graphene: hybrid materials for energy storage. *Angew. Chem.* **2018**, *130*, 1046–1050.

(79) Lukose, B.; Kuc, A.; Heine, T. The structure of layered covalent-organic frameworks. *Chem. Eur. J.* **2011**, *17*, 2388–2392.

(80) Lukose, B.; Kuc, A.; Frenzel, J.; Heine, T. On the reticular construction concept of covalent organic frameworks. *Beilstein J. Nanotechnol.* **2010**, *1*, 60–70.

(81) Kuc, A.; Springer, M. A.; Batra, K.; Juarez-Mosqueda, R.; Wöll, C.; Heine, T. Proximity Effect in Crystalline Framework Materials: Stacking-Induced Functionality in MOFs and COFs. *Adv. Funct. Mater.* **2020**, *30*, 1908004.

(82) Thomas, M.; Brehm, M.; Kirchner, B. Voronoi dipole moments for the simulation of bulk phase vibrational spectra. *Phys. Chem. Chem. Phys.* **2015**, *17*, 3207–3213.

(83) Heyden, M.; Sun, J.; Funkner, S.; Mathias, G.; Forbert, H.; Havenith, M.; Marx, D. Dissecting the THz spectrum of liquid water from first principles via correlations in time and space. *Proc. Am. Acad. Arts Sci.* **2010**, *107*, 12068–12073.

(84) Zhang, X.; Han, W.; Wu, J.; Milana, S.; Lu, Y.; Li, Q.; Ferrari, A.; Tan, P. Raman spectroscopy of shear and layer breathing modes in multilayer MoS_2. *Phys. Rev. B* **2013**, *87*, 115413.

(85) Silvestrelli, P. L.; Parrinello, M. Structural, electronic, and bonding properties of liquid water from first principles. *J. Chem. Phys.* **1999**, *111*, 3572–3580.

(86) Silvestrelli, P. L.; Parrinello, M. Water molecule dipole in the gas and in the liquid phase. *Phys. Rev. Lett.* **1999**, *82*, 3308.

(87) Stenger, F. Summary of Sinc numerical methods. *J. Comput. Appl. Math.* **2000**, *121*, 379–420.

(88) Frigo, M.; Johnson, S. G. FFTW: An adaptive software architecture for the FFT. Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181). 1998; pp 1381–1384.

(89) Brehm, M.; Kirchner, B. TRAVIS - A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories. *J. Chem. Inf. Model.* **2011**, *51*, 2007–2023.

(90) Brehm, M.; Thomas, M.; Gehrke, S.; Kirchner, B. TRAVIS—A free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* **2020**, *152*, 164105.

(91) Knoop, F.; Purcell, T. A. R.; Scheffler, M.; Carbogno, C. FHI-vibes: Ab Initio Vibrational Simulations. *J. Open Source Softw.* **2020**, *5*, 2671.

(92) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.

(93) Vanden-Eijnden, E.; Ciccotti, G. Second-order integrators for Langevin equations with holonomic constraints. *Chem. Phys. Lett.* **2006**, *429*, 310–316.

(94) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(95) Tkatchenko, A.; Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, *102*, 073005.

(96) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164.

(97) Kühne, T. D. et al. CP2K: An electronic structure and molecular dynamics software package-Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, 194103.

(98) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703.

(99) Hartwigsen, C.; Gœdecker, S.; Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641.

(100) Krack, M. Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor. Chem. Acc.* **2005**, *114*, 145–152.

(101) VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **2007**, *127*, 114105.

(102) Bas, E. E.; Garcia Alvarez, K. M.; Schneemann, A.; Heine, T.; Golze, D. Robust Computation and Analysis of Vibrational Spectra of Layered Framework Materials including Host-Guest Interactions. **2024**, 10.5281/zenodo.11065943.

(103) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

(104) Hermann, J.; Tkatchenko, A. Density functional model for van der Waals interactions: Unifying many-body atomic approaches with nonlocal functionals. *Phys. Rev. Lett.* **2020**, *124*, 146401.

(105) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

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

(107) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(108) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.

(109) Sabatini, R.; Gorni, T.; De Gironcoli, S. Nonlocal van der Waals density functional made simple and efficient. *Phys. Rev. B* **2013**, *87*, 041108.

(110) Peng, H.; Yang, Z.-H.; Perdew, J. P.; Sun, J. Versatile van der Waals density functional based on a meta-generalized gradient approximation. *Phys. Rev. X.* **2016**, *6*, 041005.

(111) Bohm, D.; Pines, D. A collective description of electron interactions: III. Coulomb interactions in a degenerate electron gas. *Phys. Rev.* **1953**, *92*, 609.

(112) Gell-Mann, M.; Brueckner, K. A. Correlation energy of an electron gas at high density. *Phys. Rev.* **1957**, *106*, 364.

(113) Liang, L.; Zhang, J.; Sumpter, B. G.; Tan, Q.-H.; Tan, P.-H.; Meunier, V. Low-frequency shear and layer-breathing modes in Raman scattering of two-dimensional materials. *ACS nano* **2017**, *11*, 11777–11802.

(114) Wang, M. et al. Unveiling electronic properties in metal–phthalocyanine-based pyrazine-linked conjugated two-dimensional covalent organic frameworks. *J. Am. Chem. Soc.* **2019**, *141*, 16810–16816.

TOC Graphic