END-TO-END DIFFERENTIABLE FORCE FIELD GENERATOR WITH CRYSTAL STRUCTURE DIFFERENTIATION AND MATCHING

A PREPRINT

Hiroshi Nakano*, **Shinnosuke, Hattori**, **Hajime Kobayashi**, **Takumi Araki**,
Masakazu Ukita, **Toshio Nishi**, **Yoshihiro Kudo**
Material and Device Analysis Center, Sony Semiconductor Solutions
4-14-1, Asahi-cho, Atsugi, Kanagawa, Japan
*Hiroshi.B.Nakano@sony.com

ABSTRACT

Differentiable programming has accelerated the development of force-field (FF) parameterization techniques. Specifically, automatic differentiation (AD) facilitates energy and force matching by differentiating them with respect to the FF parameters; hereinafter, referred to as force differentiation and matching (FDM). Conversely, crystal structure matching with AD has persisted as a challenge because the converged structures optimized by the iterative algorithm cannot be differentiated with respect to the FF parameters. Therefore, in this paper, we propose a structure differentiation and matching (SDM) method, wherein the converged structures are directly differentiated using the parameters with implicit function differentiation and matched with the experimental crystal structures. Subsequently, with a case study, we compared the reproducibility of the crystal structures, internal atomic coordinates, and lattice energies on eight exemplary molecules with the differentiable Ewald method for long-range interactions. The results indicated that SDM outperformed FDM on all three criteria. The FFs generated by SDM reproduced the lattice constants with a mean error of 0.56 %, the internal atomic coordinates with an error of 0.16 Å, and the lattice energies with an error of 0.14 kcal/mol. The corresponding accuracies obtained with FDM were 1.2 %, 0.22 Å, and 2.40 kcal/mol, respectively. Furthermore, we performed molecular dynamics simulations on a supercell, containing more than 3000 atoms, to confirm if the crystal structures were preserved under temperature fluctuations at 300 K. Overall, this method is not limited to Amber-type FFs and can be easily applied to the other types of FFs. Thus, we believe that SDM will emerge as one of the new standards for parameterizing FFs with crystal structures.

Keywords Force field, differentiable programming, implicit function differentiation, automatic differentiation, organic molecule, crystal structure

1 Introduction

Molecular dynamics (MD) is an essential tool for atomic and molecular modeling of small organic molecules [1–3]. In structure-based drug design, the characterization of intermolecular interactions is crucial for predicting ligand activity [4–7]. MD provides vital insights into the fundamental mechanisms governing the condensed matter properties of a diverse range of materials [8–10]. Moreover, in studies of quantum mechanics (QM) and molecular mechanics (MM), MD is required to create a realistic MM region [11–19].

In particular, the accuracy of MD relies on three key factors: type of force fields (FFs), reference data, and FF optimization algorithms. Specifically, FFs can be classified into classical FFs, *e.g.*, generated amber force field (GAFF) [3, 20–22], and neural network potential (NNP) [23–39]. Although NNPs have garnered significant attention in recent years, classical FFs remain highly interpretable and beneficial owing to its simple equations and fewer parameters [3, 20–22]. Thus, this study was focused on the generation of classical FFs.

The reference data can be either QM calculations, or experimental data [36]. In case of experimental data, an active research area includes the FF optimization techniques to reproduce measurable physical properties [36, 40, 41]. In this regard, the crystal structure obtained from single-crystal X-ray diffraction forms one of the most vital experimental datasets. This is important for small organic molecules, because it contains intermolecular structure data. The reproducibility of crystal structures has been generally used to validate FFs [42–44]. , whereas the cohesive properties (*e.g.*, lattice energy) have been considered to discuss the stability or strength of the intermolecular interactions [45–47]. Categorically, these crystal structures have been compiled in Cambridge Structural Database (CSD), and the data on sublimation enthalpies—convertible to lattice energies—is readily accessible as well [48]. Thus, we focus on the development of FF generation techniques that can reproduce crystal structures and lattice energies.

The FF optimization algorithms can be classified into two categories: non-differentiating and differentiating methods. In particular, non-differentiating algorithms include Bayesian optimization methods [49–53], evolutionary algorithms [54], and particle swarm optimization [55, 56], whereas the differentiating methods apply an optimization algorithm containing derivative coefficients to differentiate a function evaluating the FFs. The coefficients enable an efficient optimization for numerous variables such as the gradient descent methods or quasi-Newton methods.

For instance, a software package, ForceBalance, has been developed to generate FF using numerical differentiation [22, 40, 41, 57–60]. More recently, certain methods have been proposed based on automatic differentiation (AD) to mitigate the approximation errors and improve the efficiency of derivative calculations [61–65]. Accordingly, in this study, we focused on AD-based FF generation techniques by utilizing the crystal structures and lattice energies as the reference data.

However, FF generation with reference to crystal structures is a challenging task, because the crystal structures do not provide information on the energies and forces acting on atoms. The only information obtainable from them is that the structure is stable compared to its surrounding structures. In contrast, the energy and force information in various structures—required as training data in NNP and conventional differentiable FF parameterizations—cannot be obtained solely from the crystal structures [23–39, 44, 65]. Although the energy and force of structures near the crystal structure can be evaluated using quantum mechanical formulations, the comprehensive calculation of forces in unstable structures becomes computationally expensive because of the highly accurate intermolecular interactions. In addition, FF optimization with AD poses certain limitations in case of using the crystal structure as the reference data. The crystal structures were evaluated using an iterative structure-optimization algorithm to converge the forces into zero. Unfavorably, the convergence structures cannot be automatically differentiated with the FF parameters, because all the iterations must be connected with the chain rule of differentiation in the AD. Thus, it forms an immensely deep network structure that requires extensive memory and renders the optimization unstable.

To circumvent this issue, a noraml approach utilizes the information that the forces acting on each atom are balanced to zero. The calculation of forces in stable structures using FFs follows a deterministic algorithm instead of a convergence one, which enables the differentiation of forces on the atoms with respect to FF parameters. By utilizing these derivative coefficients, the optimization algorithms can be applied using gradients, referred to as force differentiation and matching (FDM).

However, the optimization goal of zero forces acting on the experimental crystal structures (*e.g.*, FDM) differs from the objective of ensuring correspondence between the simulated and experimental structures. For optimization algorithms, this implies the variations among the evaluation functions.

Here, we propose a method based on implicit function differentiation (IFD) to differentiate the simulated crystal structures with respect to the FF parameters [66].

Using the derivative coefficients, we can optimize the FF parameters to ensure correspondence between the converged and experimental structures. This method is referred to as structure differentiation and matching (SDM) that developed an automatic FF generation program with the reference data: (i) stable single-molecule structure, (ii) crystal structure, (iii) crystallization energy, and (iv) Potential Energy Surfaces (PESs) of free rotating dihedral angles. In the case study, we compare SDM with FDM with eight materials including anthracene, biphenyl (BIPHEN), and benzoic acid (BENZAC). In addition, we optimized the charges considering the long-range Coulomb interactions in crystals. To accurately represent the potential, general MDs employ the Ewald method [67], or its faster version, and the particle mesh Ewald (PME) method [68]. The optimization problem of FF generation is complicated by the need to optimize the charge of each atom as a variable to account for the long-range interactions with several charges located outside the unit cell. Thus, we implemented the Ewald method in its differentiable form.

Notably, the experimentally derived crystal structure in CSD with single-crystal X-ray diffraction may differ from the true stable structure, as it must be theoretically measured at absolute zero. In this study, we considered negligible deviations between the experimental and stable structures.

The optimized parameters can be obtained as an output corresponding with the LAMMPS form [2] which are compatible because the energies evaluated by the program and LAMMPS are consistent, with errors less than 1.0×10^{-3} kcal/mol. Therefore, the FFs generated by the proposed program are readily applicable to large-scale MD simulations with LAMMPS.

2 Methods

2.1 Force Field

We used Amber-type FFs to ensure compatibility with the existing MD softwares [1, 2], The FFs are expressed as

$$E_{\text{total}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihed}} + E_{\text{vdw}} + E_{\text{coul}}, \tag{1}$$

$$E_{\text{bond}} = \sum_{\text{bond } r} K_r (r - r_{\text{eq}})^2, \tag{2}$$

$$E_{\text{angle}} = \sum_{\text{angle } \theta} K_\theta (\theta - \theta_{\text{eq}})^2 \ , \tag{3}$$

$$E_{\text{dihed}} = \sum_{\text{dihed } \phi} \sum_{n=1}^{4} \frac{V_n}{2} [1 + \cos(n\phi - \gamma_n)] \ , \tag{4}$$

$$E_{\text{vdW}} = \sum_{i<j} 4\epsilon \left[(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^6 \right] \ , \tag{5}$$

$$E_{\text{coulomb}} = \sum_{i<j} \frac{q_i q_j}{\epsilon_p r_{ij}}. \tag{6}$$

The total energy of the system E_{total} comprises bond energy E_{bond}, angular energy E_{angle}, dihedral angular energy E_{dihed}, van der Waals (vdW) force potential E_{vdw} in the Lenard-Jones (LJ) form and Coulomb potential E_{coul}. The variables to be optimized for constructing the FFs are shown in Table 1. A vector with all variables defined in Table 1 as elements is denoted by **p**. The γ_n of the dihedral angle was set to 0 or π, because the same effect can be obtained by optimizing V_n. We adopted the coefficients used in OpenFF [22]. The coefficient of charge q_i was set to 0.1 such that the variation is comparable to that of other variables. Furthermore, a constraint condition was introduced such that the sum of the charges was zero for each molecule.

In particular, we used as many types of vdW parameters and atomic charges as possible, considering the symmetry of the molecule, because the optimization efficiency through the differentiation methods will not deteriorate with the increasing number of variables compared to the non-differentiation methods [49–56]. As an exemplary case, benzoic acid (BENZAC) is presented in Figure 1. In total, we selected 11 types of atoms—C1 to C5, H1 to H4, O1, and

Table 1: Regularization factor

Type	Regularization factor
bond force constant K_r	100 kcal/mol/Å^2
equilibrium bond length r_{eq}	0.1 Å
angle force constant K_θ	100 kcal/mol^{-1} rad^2
equilibrium angle θ_{eq}	20 degrees
dihedral force constant V_n	1 kcal/mol
vdW well depth ϵ	0.1 kcal/mol
vdW lendth σ	1 Å
charge q	0.1 e

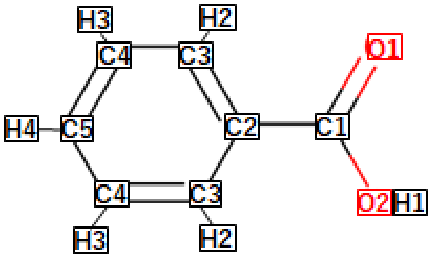


Figure 1: Symmetry-based atom type assignment

O2—and generated 22 vdW parameters and 11 atomic charges, considering the symmetry of the molecular structure, thereby providing a higher degree of freedom to the FFs. This is based on the concept of equivariance that generally denotes the translational and rotational symmetry, and this concept was extended to equivalent atoms on the molecular graph. The LJ potentials of various types of atom can be evaluated using arithmetic mixing [1, 2, 21].

In case of calculating the crystal structures with periodic structures in MD, the long-range interactions caused by the Coulomb potential must be considered. The Ewald method was implemented in a differentiable form. The PME method is commonly applied because it approximates the Ewald method with a lower computational cost. Nonetheless, our aim was to propose a new FF construction method reproducing the crystal structures of a small molecule. The FFs were primarily evaluated for a unit lattice, and large systems were not investigated. This restricts the merit of PME to small-sized molecules. Therefore, we applied the theoretically simpler Ewald method, which is expressed as

$$E_{\mathrm{coulomb}} = E^S + E^L - E^{\mathrm{self}} \tag{7}$$

$$E^S = \frac{1}{2\epsilon_p}\sum_{\mathbf{n}}\sum_{i=1}^{N}\sum_{j=1}^{N}{}' \frac{q_i q_j}{|r_{ij}+\mathbf{n}\cdot\mathbf{L}|}\mathrm{erfc}\left(\frac{|\mathrm{r_{ij}}+\mathbf{n}\cdot\mathbf{L}|}{\sqrt{2}\sigma}\right) \tag{8}$$

$$E^L = \frac{2\pi}{V\epsilon_p}\sum_{k\neq 0}\frac{\mathrm{e}^{-\sigma^2\mathrm{k}^2/2}}{k^2}|S(\mathbf{k})|^2 \tag{9}$$

$$E^{\mathrm{self}} = \frac{1}{\epsilon_p}\frac{1}{\sqrt{2\pi}\sigma}\sum_{i=1}^{N} q_i^2. \tag{10}$$

To ensure compatibility with LAMMPS, we implemented the neighbor lists and special bonds. Furthermore, a CHARMM-style energy-switching function was applied to E_{vdW} [69]. Expressed as a LAMMPS keyword, the program implemented "lj/charmm/coul/long"-type non-bonding interactions. The remaining parameters (dielectric constant ϵ_p and distance threshold to account for Coulomb potential and LJ potential for atoms within three topological distances (special bonds)) were considered to be constants. The dielectric permittivity ϵ_p was 3.0, the special bonds was set to the DREIDING type, and the parameters for the improper type of dihedral angles were set to the same values as those in GAFF. As detailed in Appendix 1 (A1), for all molecules tested herein, the energies calculated by the proposed program were consistent with LAMMPS, with errors less than 1.0×10^{-3}. Furthermore, the program exhibits the capability to output the FFs in the LAMMPS format.

2.2 Evaluation Function

In this study, the evaluation functions for FF optimization comprised the (i) evaluation function L^{M} for the monomer structure, (ii) crystal structure evaluation function L^{C}, (iii) evaluation function of the lattice energy L^{E}, and (iv) PES evaluation functionL_l^{P} for dihedral angle l of freely rotating bonds. As more than one dihedral angles can exist in a molecule, the evaluation function for PESs can be expressed as the summation of the corresponding terms. The evaluation function is defined as follows.

$$L^{\mathrm{all}} = w^{\mathrm{M}} L^{\mathrm{M}} + w^{\mathrm{C}} L^{\mathrm{C}} + w^{\mathrm{E}} L^{\mathrm{E}} + \sum_{l} w_l^{\mathrm{P}} L_l^{\mathrm{P}}, \tag{11}$$

where $w^{\mathrm{M}}, w^{\mathrm{C}}, w^{\mathrm{E}}$ and w_l^{P} denote the weights balancing the corresponding terms. To increase the accuracy of a particular evaluation item, these weights can be adjusted based on the research objective. In particular, these weights including the values and its backgrounds are elaborated in Appendix 2 (A2). In addition, we introduced a constraint that the sum of the charges of the atoms in a molecule equals to the total charge Q. As all the selected molecules were neutral, Q was set to zero. The constraints defined by the inequalities were introduced to limit the search ranges to physically reasonable values, as detailed in Appendix 3 (A3). For the conditions of charge constraints and variables ranges, we adopted the optimization method of sequential least-squares programming (SLSQP) that can optimize the FF parameters with the constraints and inequality relations.

2.2.1 Evaluation function for single molecule structures

The four terms in Eq. (11) are explained as follows. First,L^{M} quantifies the deviation between the reference single-molecule structure and the converged structure optimized with a FF. The structure was optimized with the gradient descent method in JAXOPT [70]. To compare the structures, we used the internal coordinates $\mathbf{u}^{\mathrm{M}} = (\mathbf{r}, \boldsymbol{\theta}, \boldsymbol{\phi})$ instead of cartesian coordinates $\mathbf{u}$, where $\mathbf{r}$ denotes the bond length, $\boldsymbol{\theta}$ indicates the bond angle, and $\boldsymbol{\phi}$ represents the dihedral angle. The internal coordinates are symmetric (equivalence) for the translational and rotational operations. Additionally, the Amber-type FF is expressed in a form that presumes the internal coordinates as arguments, thereby rendering the internal coordinates suitable for evaluating the structural similarity of the individual molecules. The reference structure is denoted as $\mathbf{u'}^{\mathrm{M}}$ and the structure optimized with FF is denoted as $\tilde{\mathbf{u}}^{\mathrm{M}}$. The evaluation function for single molecule structures can be expressed as

$$L^{\mathrm{M}} = (\tilde{\mathbf{u}}^{\mathrm{M}}(\mathbf{p}) - \mathbf{u'}^{\mathrm{M}})^2. \tag{12}$$

where $\mathbf{r}$, θ, and ϕ indicate the elements in the same vector $\mathbf{u'}^{\mathrm{M}}$, despite bearing distinct units. The regularization coefficients in Table 1 were multiplied with the value of $\mathbf{r}$, θ, and ϕ respectively, to ensure correspondence with the same vector $\mathbf{u'}^{\mathrm{M}}$. For instance, a bond length variation of 0.1 Å and an angular difference of 20° pose the same impact on evaluation function L^{M} [22].

2.2.2 Evaluation function for crystal structures

For crystal structure matching, the lattice vectors matching as well as the atomic coordinates are required. The evaluation function are defined as

$$L^{\mathrm{C}} = (\tilde{\mathbf{u}} - \mathbf{u'})^2 + c_L(\tilde{T}_v - T'_v)^2. \tag{13}$$

The coefficient of cL was set to 10, because the variations in the lattice vector Tv influences all atoms in the crystal and should be regarded as more essential than a coordinate of an atom.

The three lattice vectors $\tilde{T}_v$ in the converged structures have to be should be fitted to the experimental structures T'_v. The equivalency for rotation and translational symmetries is not necessary in the crystal structure matching, because the crystal lattices induce a loss of symmetries. Thus, the similarity of the crystal structures was compared in terms of cartesian coordinates $\mathbf{u}$. As stated herein, we used the extended coordinates $\tilde{\mathbf{u}}^{\mathrm{C}} = (\tilde{\mathbf{u}}^T, T_v^T)^T$, where ^{T} denotes transpose. The coefficient of c_L was set to be 10, because the variations in the lattice vector T_v influences all atoms in the crystal and should be regarded as more essential than a coordinate of an atom.

2.2.3 Evaluation function for lattice energy

The lattice energy U'_{lat} was computed from the experimental accessible values of the sublimation enthalpy H_{subl} that is correlated with the lattice energy,

$$H_{\mathrm{subl}} = -U'_{\mathrm{lat}} - 2RT \tag{14}$$

where R denotes the gas constant and T indicates the temperature [48]. Based on the crystal structure energy per unit cell E^{C} and the single molecule energy E^{M}, the lattice energy $\tilde{U}_{\mathrm{lat}}$ can also be defined as

$$\tilde{U}_{\mathrm{lat}} = E^{\mathrm{M}} - \frac{E^{\mathrm{C}}}{N_{\mathrm{mol}}}, \tag{15}$$

The evaluation function for the lattice energy matching L^{E} can be evaluated as

$$L^{\mathrm{E}} = (U^{\mathrm{lat}} - \tilde{U}^{\mathrm{lat}}(\tilde{\mathbf{u}}^{\mathbf{C}}, \mathbf{p}))^2. \tag{16}$$

An FF reproducing both the crystal structure and U^{lat}, enables the simulation of MD with the FF to reflect the stability of the crystal [48].

2.2.4 Evaluation function of PESs

The total evaluation function should include the evaluation functions for matching PES of freely rotatable dihedral angles. The reference data were acquired with the QM calculation of a single molecule. The evaluation functions of the energy surface PES_l along with the rotation of the dihedral angles l can be defined as

$$L_l^{\mathrm{P}} = \sum_{\mathbf{u} \in \mathrm{PES}_l} \exp(-\min(E(\mathbf{u}, \mathbf{p}), E'(\mathbf{u}))/2k_B T_{\mathrm{PES}})(E(\mathbf{u}, \mathbf{p}) - E'(\mathbf{u}))^2, \tag{17}$$

where the term $\exp(-\min(E(\mathbf{u}, \mathbf{p}), E'(\mathbf{u}))/2k_B T_{\mathrm{PES}})$ includes the coefficients for prioritizing the points near the lowest points in the PES. k_B denotes the Boltzmann constant, and T_{PES} represents the virtual temperature for prioritization, which was set to 2000 K according to Ref. [22]. The weight w^{E} was set to $0.25/N_{\mathrm{atom}}/N_{\mathrm{points}}$ (kcal/mol)$^{-2}$, where N_{points} denotes the number of points in a PES.

2.3 FF optimization methods

2.3.1 Stable structure Differentiation (SDM)

SDM computes the derivative coefficients of the evaluation function Eq.(11) with respect to the FF parameters by employing direct differentiation of the stable structures computed with an iterative optimization algorithm. Figure 2 illustrates the block diagram of SDM. The block diagram of SDM is presented in Figure 2, wherein the energy of the systems evaluated by substituting $\tilde{\mathbf{u}}$ and $\mathbf{p}$ into an Amber-type energy function formula (Eq. (1)) is presented in Figure 2(a). The block diagram for the differentiations of the stable structures with respect to $\mathbf{p}$ is illustrated in Figure 2(b). In the top block, the iterative process was followed to obtain $\tilde{\mathbf{u}}$ using the block defined in Figure 2(a). In conventional implementations, AD was repeated with the same number of iterations of the convergence. Despite restricting to the short-range interaction E^S, almost 10,000 varying Coulomb potentials existed between the two atoms in each layer, which were registered in the neighborhood list. Therefore, if 100 cycles are required for structural structure optimization, it would constitute an enormous NN comprising at least 1 million neurons.

On the contrary, the following SDM enabled the efficient computation of the derivative. First, the converged structure $\tilde{\mathbf{u}}$ satisfies

$$\mathbf{F}(\tilde{\mathbf{u}}(\mathbf{p}), \mathbf{p}) \equiv \frac{\partial E(\tilde{\mathbf{u}}(\mathbf{p}), \mathbf{p})}{\partial \mathbf{u}} = 0, \tag{18}$$

Subsequently, both sides of Eq. (18) were differentiated with respect to $\mathbf{p}$ to obtain

$$\frac{\partial \mathbf{F}(\tilde{\mathbf{u}}, \mathbf{p})}{\partial \mathbf{p}} = \frac{\partial \mathbf{F}(\tilde{\mathbf{u}}, \mathbf{p})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{p}} + \frac{\partial \mathbf{F}(\tilde{\mathbf{u}}, \mathbf{p})}{\partial \mathbf{p}} = 0. \tag{19}$$

Thus,

$$\frac{\partial^2 E(\tilde{\mathbf{u}}, \mathbf{p})}{\partial \mathbf{u}^2} \frac{\partial \mathbf{u}}{\partial \mathbf{p}} + \frac{\partial^2 E(\tilde{\mathbf{u}}, \mathbf{p})}{\partial \mathbf{u} \partial \mathbf{p}} = 0, \tag{20}$$

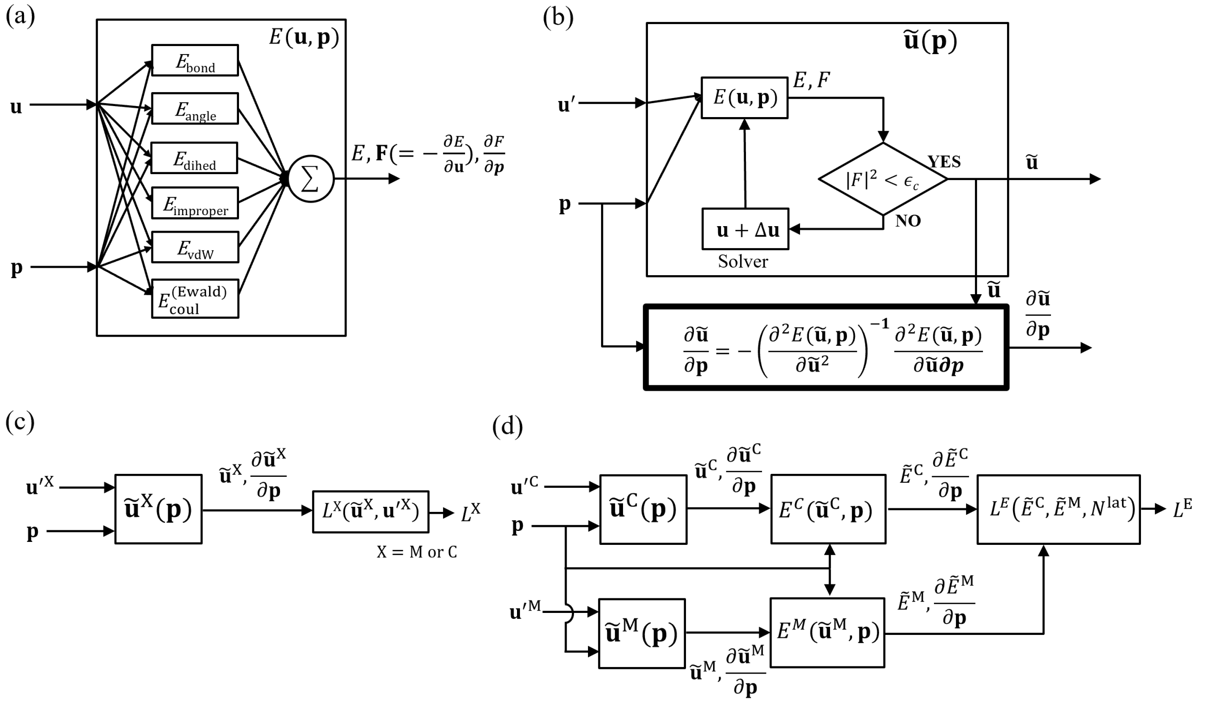

(a)
$E(\mathbf{u},\mathbf{p})$
$\mathbf{u}$
$\mathbf{p}$
E_{bond}
E_{angle}
E_{dihed}
E_{improper}
E_{vdW}
$E_{\text{coul}}^{(\text{Ewald})}$
$\sum$
$E, \mathbf{F}(= -\frac{\partial E}{\partial \mathbf{u}}), \frac{\partial F}{\partial p}$
(b)
$\tilde{\mathbf{u}}(\mathbf{p})$
$\mathbf{u}'$
$\mathbf{p}$
$E(\mathbf{u},\mathbf{p})$
E, F
$|F|^2 < \epsilon_c$
YES
NO
$\mathbf{u} + \Delta\mathbf{u}$
Solver
$\tilde{\mathbf{u}}$
$\tilde{\mathbf{u}}$
$\frac{\partial \tilde{\mathbf{u}}}{\partial \mathbf{p}}$
$\frac{\partial \tilde{\mathbf{u}}}{\partial \mathbf{p}} = -\left(\frac{\partial^2 E(\tilde{\mathbf{u}},\mathbf{p})}{\partial \tilde{\mathbf{u}}^2}\right)^{-1} \frac{\partial^2 E(\tilde{\mathbf{u}},\mathbf{p})}{\partial \tilde{\mathbf{u}} \partial \boldsymbol{p}}$
(c)
$\mathbf{u}'^{\text{X}}$
$\mathbf{p}$
$\tilde{\mathbf{u}}^{\text{X}}(\mathbf{p})$
$\tilde{\mathbf{u}}^{\text{X}}, \frac{\partial \tilde{\mathbf{u}}^{\text{X}}}{\partial \mathbf{p}}$
$L^{\text{X}}(\tilde{\mathbf{u}}^{\text{X}}, \mathbf{u}'^{\text{X}})$
L^{X}
X = M or C
(d)
$\mathbf{u}'^{\text{C}}$
$\mathbf{p}$
$\mathbf{u}'^{\text{M}}$
$\tilde{\mathbf{u}}^{\text{C}}(\mathbf{p})$
$\tilde{\mathbf{u}}^{\text{C}}, \frac{\partial \tilde{\mathbf{u}}^{\text{C}}}{\partial \mathbf{p}}$
$E^{C}(\tilde{\mathbf{u}}^{\text{C}}, \mathbf{p})$
$\tilde{E}^{\text{C}}, \frac{\partial \tilde{E}^{\text{C}}}{\partial \mathbf{p}}$
$L^{E}(\tilde{E}^{\text{C}}, \tilde{E}^{\text{M}}, N^{\text{lat}})$
L^{E}
$\tilde{\mathbf{u}}^{\text{M}}(\mathbf{p})$
$\tilde{\mathbf{u}}^{\text{M}}, \frac{\partial \tilde{\mathbf{u}}^{\text{M}}}{\partial \mathbf{p}}$
$E^{M}(\tilde{\mathbf{u}}^{\text{M}}, \mathbf{p})$
$\tilde{E}^{\text{M}}, \frac{\partial \tilde{E}^{\text{M}}}{\partial \mathbf{p}}$

Figure 2: Block diagram of (a) SDM, (b) FDM. AD of the convergence algorithm with IFT in SDM. $E(\mathbf{u}, \mathbf{p})$ in (a) is included in diagram (b). ϵ_c denotes convergence criterion, and $\Delta\mathbf{u}$ indicates the updates of $\tilde{\mathbf{u}}$. In the bottom block with bold lines, derivative coefficient $\partial\tilde{\mathbf{u}}/\partial\mathbf{p}$ can be computed using the converged structure $\tilde{\mathbf{u}}$ optimized in the top block. (c) Block diagram for partial evaluation function of $L^{\text{X}}(X = M, C)$. (d) Block diagram for partial evaluation function of L^{E}. Differentiation of function requires two derivative coefficients of $\tilde{\mathbf{u}}^{\mathbf{M}}(\mathbf{p})$ and $\tilde{\mathbf{u}}^{\mathbf{C}}(\mathbf{p})$, which are derived with IFT.

where $\partial^2 E(\tilde{\mathbf{u}}, \mathbf{p})/\partial\tilde{\mathbf{u}}^2$ epresents the Hessian defined as the second-order differentiation based on the atomic coordinates $\tilde{\mathbf{u}}$. The multiplication of the inverse matrix to both sides of Eq. (20) yields

$$\frac{\partial \tilde{\mathbf{u}}}{\partial \mathbf{p}} = -\frac{\partial^2 E(\tilde{\mathbf{u}}, \mathbf{p})}{\partial \mathbf{u}^2}^{-1} \frac{\partial^2 E(\tilde{\mathbf{u}}, \mathbf{p})}{\partial \mathbf{u} \partial \mathbf{p}}. \tag{21}$$

Although $\tilde{\mathbf{u}}$ can be calculated following the iterative process of the convergence algorithm, the differentiation of $\tilde{\mathbf{u}}$ does not have to be an iterative process, as expressed in Eq. (21). The top block computes the structure optimization using the iterative process to derive $\tilde{\mathbf{u}}$, based on the block diagram depicted in Figure 2(a). The bottom block performs the differentiation with respect to $\tilde{\mathbf{u}}$ from the top block in a non-iterative process. In Eq. (21), the second term refers to the differentiation of $\mathbf{F}$ with respect to $\mathbf{p}$ at $u = \tilde{\mathbf{u}}$, and the first term replaces the derivative variable from $\mathbf{p}$ to $\mathbf{u}$. The first term exhibits no considerable differences in the computation costs. The inverse matrix calculation poses no significant impact. Overall, the IFD facilitates the computation of $\partial\tilde{\mathbf{u}}/\partial\mathbf{p}$ with the same order of the cost of a single $\partial\mathbf{F}/\partial\mathbf{p}$ calculation. As depicted in Figure 2(c), $\partial\tilde{\mathbf{u}}/\partial\mathbf{p}$ is embedded in the calculation and differentiation of $L^{\text{X}}(\text{X} = \text{M}, \text{C})$, and as displayed in Figure 2(d), L^{E} can be explained based on the derivatives.

2.3.2 Force Differentiation and Matching (FDM)

FDM optimizes an FF by differentiating the force on the atoms with respect to $\mathbf{p}$. The method of matching the energies and forces obtained from the QM calculation is commonly used in NNP [23–39, 65].

Although, previous studies have considered several types of energies and forces for non-stable structures, the data from these reference structures cannot be used, except those related to the stable structures. As FDM can utilize only the condition $\mathbf{F}(\mathbf{u}', \mathbf{p}) = 0$ at the reference structure $\mathbf{u}'$, it replaces L^{M}, L^{C}, and L^{C} to derive into

$$L^{\mathrm{M,FDM}} = |\mathbf{F}^{\mathrm{M}}(\mathbf{u}', \mathbf{p})|^2 \tag{22}$$

$$L^{\mathrm{C,FDM}} = |\mathbf{F}^{\mathrm{C}}(\mathbf{u}', \mathbf{p})|^2 \tag{23}$$

$$L^{\mathrm{E,FDM}} = (U^{\mathrm{lat}} - \tilde{U}^{\mathrm{lat}}(\mathbf{u}'^{C}, \mathbf{p}))^2, \tag{24}$$

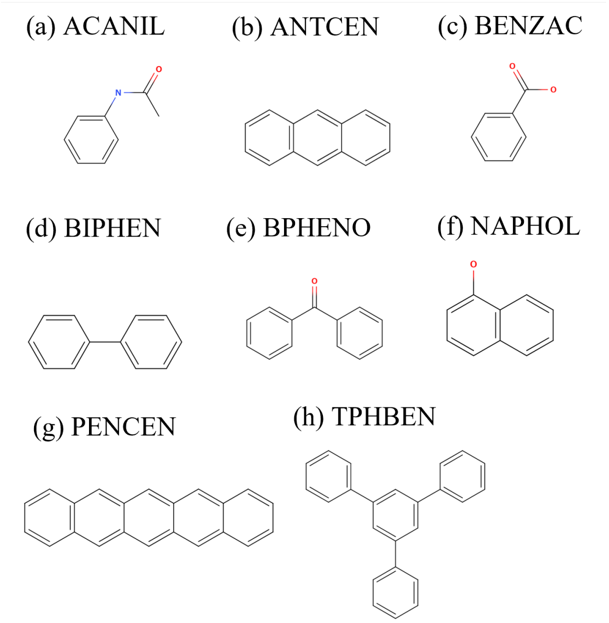
respectively. $L^{\mathrm{M,FDM}}$ and $L^{\mathrm{C,FDM}}$ reduce the forces on an atom at the reference data as zero. The crystal structure $U^{\mathrm{lat}}(\mathbf{u}'^{\mathrm{C}}, \mathbf{p})$ is used in Eq. (24) because FDM cannot differentiate any functions including the converged structures $\tilde{\mathbf{u}}^{\mathrm{C}}$. In summary, the evaluation function of FDM is defined as

$$L^{\mathrm{all}}_{\mathrm{FDM}} = w^{\mathrm{M,FDM}} L^{\mathrm{M,FDM}} + w^{\mathrm{C,FDM}} L^{\mathrm{C,FDM}} + w^{\mathrm{E}} L^{\mathrm{E,FDM}} + \sum_{l} w_l^{\mathrm{P}} L_l^{\mathrm{P}}. \tag{25}$$

2.4 Details on Comparative Study of Optimization Methods

The organic molecules that we will be treating considered in this study in this study are depicted in fFigure 3. The crystal structures of these molecules are documented in the CSD, and the Hsubl data is are reported in previous literature [48]. The selection of the eight molecules was based on their molecular weights and diversity of the substituents. Fundamentally, Tthey are composed of π-conjugated systems, such as ANTCEN and PENCEN, those with rotatable bonds, such as *e.g.*, BIPHEN and TPHBEN, and other molecules that possessing hydroxyl, carbonyl, and amino groups.

The reference structure of single molecules and PESs were calculated evaluated using by QM methods on Gaussian 16, applied with the density functional theory. The basis set employed was 6-311++g(d,p), and the functional ωB97XD [71]. The PESs of the dihedral angles were created for all of the freely rotatable single bonds by evaluating the energies of the rotations of the single bonds in the molecule by 5°. The methyl group was not included excluded in this study because owing to its weak impact effect on the quality of FFs is small. The initial parameters of FF were taken sourced from GAFF [21]. For the parameters required by the Ewald methods, we used the values determined with LAMMPS, with the accuracy set to 1×10^{-4}. The FF data generation software Antechamber and Moltemplate were used to generate the initial FF parameters GAFF calculated with atomic charges from electrostatic potentials using a grid-based method (CHELG) [2, 21, 72, 73] were used The maximum number of iterations for SLSQP for FDM and SDM was 100, where each iteration includes approximately five evaluation function calculations. All methods were coded on Python 3.8.13 and JAX 0.3.13 [62]. The FF parameter optimizations were performed with SLSQP algorithms of



(a) ACANIL (b) ANTCEN (c) BENZAC

(d) BIPHEN (e) BPHENO (f) NAPHOL

(g) PENCEN (h) TPHBEN

Figure 3: Eight molecules used for evaluation. The six-letter symbols represent CSD identification IDs, and those in parentheses are generic names. (a) ACANIL, (b) ANTCEN (anthracene), (C) BENZAC (benzoic acid), (d) BIPHEN (biphenyl), (e) BPHENO (benzophenone), (f) NAPHOL (1-naphthol) (g) PENCEN (pentacene), and (h)TPHBEN (1,3,5-triphenyl benzene)

SciPy v1.8.1 [74], and the structure optimizations within the SDM algorithm were executed on JAXOPT0.5 [66]. For instance, on a computer with an Intel Xeon Gold 6230 2.10 GHz CPU and an NVIDIA Georce 2080 Ti GPU, the total optimization duration of ACANIL with FDM and SDM was 15 and 23 min, respectively.

3 Result and Discussion

3.1 Optimized Evaluation Functions

First, we evaluate the performance of the optimization techniques by assessing the evaluation functions. The values of the evaluation function L^{all} after optimization with FDM and SDM are listed in Table 2. In case of FDM, the objective function to be minimized is represented by Eq. (25), whereas the values were calculated using Eq. (11). The results

Table 2: Values of evaluation functions for GAFF (initial values) and after optimization by FDM, and SDM

CSD ID	GAFF	FDM	SDM
ACANIL	4.457	0.991	0.177
ANTCEN	0.308	0.073	0.041
BENZAC	0.890	0.476	0.026
BIPHEN	0.274	0.208	0.025
BPHENO	1.181	0.347	0.052
NAPHOL	0.445	0.478	0.063
PENCEN	0.221	0.221	0.072
TPHBEN	0.244	0.085	0.039

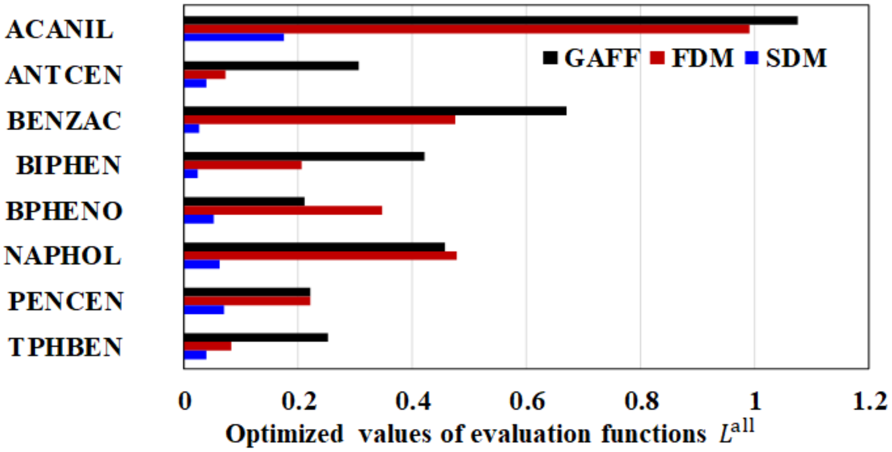


Figure 4: Bar chart of evaluation functions for GAFF (initial values) and after optimization by FDM, and SDM

of the minimization via FDM and SDM are depicted in Figure 4 as a bar chart in conjunction with the GAFF method.

For all molecules, the SDM yielded smaller values than FDM. On average, the evaluation functions were reduced to 1/6.6. For BIPHEN, the results varied by a factor of 18. Although the FDM was improved over the GAFF for all molecules, except for PENCEN, it was less accurate than the SDM. This is because the evaluation function of FDM, *i.e.*, $L^{\mathrm{all}}_{\mathrm{FDM}}$ differed from the function L^{all}. Thus, the SDM was more accurate than FDM. Moreover, the L^{all} function was used as the evaluation function for optimization and was directly differentiated with respect to the FF parameter $\mathbf{p}$ in SDM. In the following subsections, we compare the accuracy of the crystal structure and lattice energy with the initial parameter GAFF and the FDM- and SDM-optimized FFs. The comparative analyses with PESs are described in Appendix 4 (A.4) because the evaluation function of PESs L^{P}_{l} do not include the iterative process to ensure no algorithmic differences between the FDM and SDM.

3.2 Crystal structures

A histogram plotting the errors of the lattice constants optimized with the initial GAFF and the FFs optimized by FDM and SDM is presented in Figure 5. The reproducibility of both the axis length and the angles between the axes improved increasingly in the following order: GAFF, FDM, and SDM. The variations among all the molecules and the six lattice constants are detailed in Appendix 5 (A5). The FDM and SDM exhibited an average error of 1.2 % and 0.56 %, respectively. The number of lattice constants with an error greater than 2 % was significantly reduced from 12 in FDM to only one in SDM.

For instance, the variations in angle β of BIPHEN are depicted in Figure 6. The β of GAFF is approximately 90°, which generated an orthorhombic lattice, whereas that of SDM was 95.0° produced a monoclinic structure in accordance with the reference structure. The error in angle β of FDM was between those of GAFF and SDM.

To quantify the accuracies of the orientation of the molecules and the internal structures of the lattices, we calculated the mean absolute error (MAE) of the internal atomic coordinates of the molecules in the unit cell, as depicted in

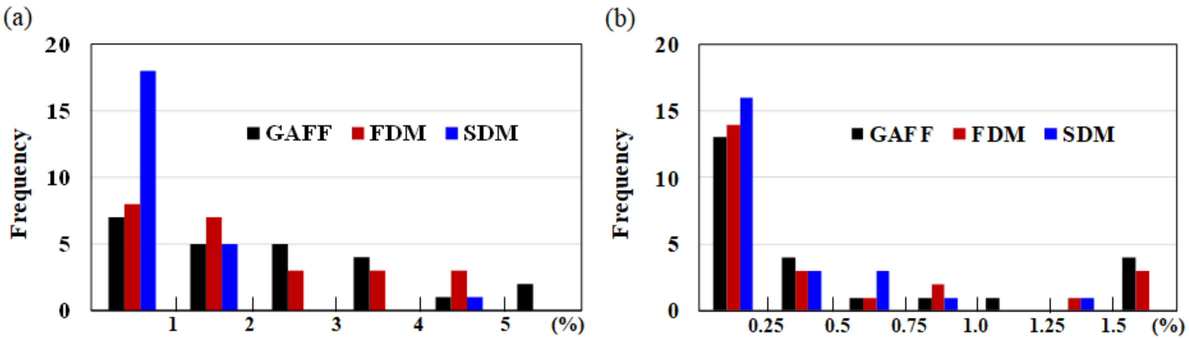


Figure 5: Histogram lattice constants optimized with GAFF, and FF derived with FDM, and SDM. (a) Histogram of error of lengths of axis: a, b, and c, wherein smallest area represents 0–1 % and largest area represents >5 %. (b) Histograms of error of angles between axes of b–c, c–a, and a–b: α, β, and γ, respectively, wherein smallest area represents 0–0.25 % and largest area represents >1.5 %.

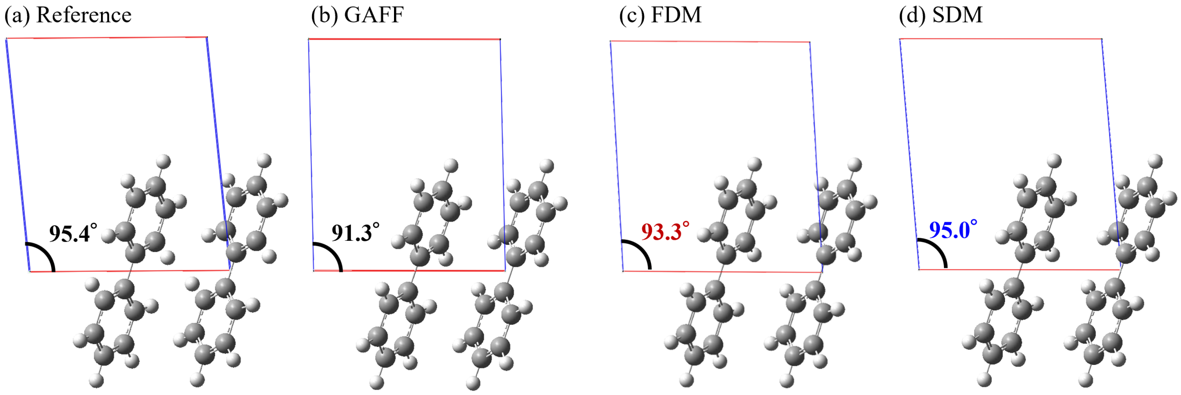

(a) Reference (b) GAFF (c) FDM (d) SDM

Figure 6: The crystal structures of BIPHEN. (a) Reference structure derived from CSD. The optimized structure (b) with GAFF, (c) the FF optimized by FDM, and (d) the FF by SDM

Figure 7. Across all molecules, the average MAE was 0.23 Å for GAFF, 0.22 Å for FDM, and 0.16 Å for SDM, which corresponds to similar improvement in the accuracy of the lattice constants.

In four molecules out of eight, FDM failed to improve the accuracy of the internal coordinates from GAFF. As depicted in Figure 8, FDM exhibited a distorted molecular structure; the red arrows indicate that the hydrogen atom attached to an aromatic carbon atom of the benzene ring is not in the plane of the benzene ring, and the planar molecule including the carboxyl group is oriented in an oblique manner. This is caused by the evaluation function for FDM, which is an overdetermined system with a smaller number of variables than the number of independent equations in Eq. (25). Consequently, it failed to yield a solution of $L^{\mathrm{all}} = 0$. Therefore, the optimized FF contained nonzero residual forces at the reference structure $\mathbf{u}'$, and the residual forces prevented the FF in ensuring that the stable structure $\tilde{\mathbf{u}}$ exists near the reference structure $\mathbf{u}'$.

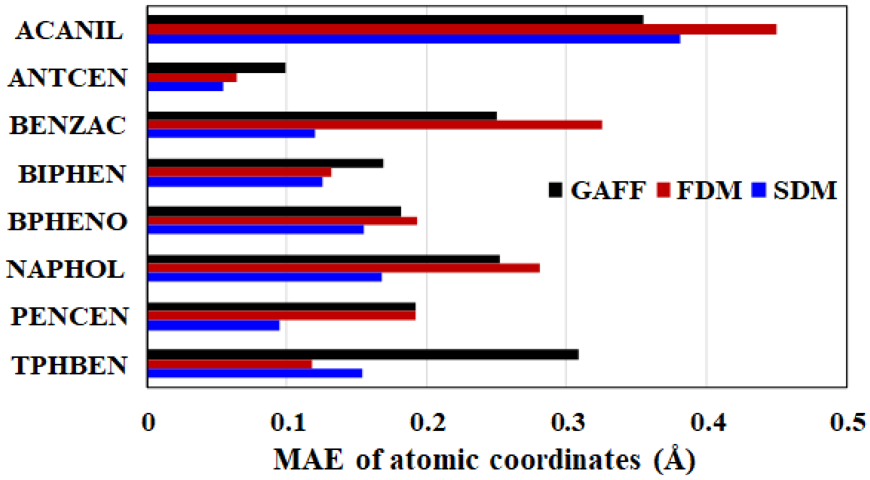


Figure 7: Crystal structures of BIPHEN. (a) Reference structure derived from CSD. Optimized structure (b) with GAFF, (c) FF optimized by FDM, and (d) FF by SDM

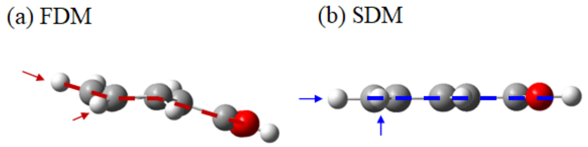


Figure 8: Optimized structures of a single BENZAC molecule by FF derived *via* (a) FDM and (b) SDM

3.3 Lattice Energies

In addition to the errors derived from the reference data, the lattice energies in GAFF, FDM, and SDM are presented in Table 3. As observed, SDM exhibited the highest accuracies. As GAFF is not specifically designed to reproduce U^{lat} for single molecules, it produced large errors. In contrast, the results of FDM for ANTCEN and TPHBEN were more consistent to the reference data than those of SDM. However, for ACANIL, NAPHOL and PENCEN, the errors were large at 6.6, 4.6, and 3.4 kcal/mol, respectively. The MAE of the lattice energies derived by FDM was 2.40 kcal/mol. Conversely, on average, SDM exhibited greater reproducibility with an MAE of 0.14 kcal/mol and the highest error of 0.34 kcal/mol for BENZAC.

Table 3: GAFF (initial values), FDM and SDM lattice energies and errors (in parentheses) from experiments (kcal/mol).

CSD ID	Reference	GAFF	FDM	SDM
ACANIL	-25.04	-19.95 (5.09)	-31.65 (-6.61)	-24.99 (0.05)
ANTCEN	-25.13	-20.46 (4.68)	-25.08 (0.05)	-25.05 (0.08)
BENZAC	-22.50	-16.27 (6.23)	-24.38 (-1.88)	-22.17 (0.34)
BIPHEN	-21.07	-17.56 (3.51)	-20.31 (0.77)	-20.97 (0.10)
BPHENO	-22.48	-19.44 (3.04)	-20.53 (1.95)	-22.59 (-0.11)
NAPHOL	-22.98	-18.41 (4.57)	-18.42 (4.56)	-22.88 (0.10)
PENCEN	-30.11	-33.52 (-3.41)	-33.52 (-3.41)	-29.98 (0.14)
TPHBEN	-36.51	-33.33 (3.18)	-36.51 (0.00)	-36.74 (-0.23)

Table 4: Lattice constants of supercell model after MD simulations at 300 K and NPT conditions and errors (in parentheses) from experimental crystal structures (kcal/mol). The lengths were divided by the multiples to effectively compare the supercells with the lattice constants of unit cells.

CSD ID	a (Å)		b (Å)		c (Å)		α (degree)		β (degree)		γ (degree)	
ACANIL	19.49	(-0.02)	9.27	(-0.09)	7.63	(-0.15)	89.82	(-0.19)	90.01	(0.01)	90.01	(0.01)
ANTCEN	16.98	(-0.15)	6.13	(0.10)	11.16	(-0.02)	90.25	(0.25)	123.95	(-0.75)	89.85	(-0.15)
BENZAC	10.87	(-0.15)	5.18	(0.02)	22.11	(0.14)	90.10	(0.10)	98.30	(0.89)	90.16	(0.16)
BIPHEN	8.18	(0.06)	5.62	(-0.02)	9.46	(-0.01)	89.88	(-0.12)	95.04	(-0.36)	90.03	(0.03)
BPHENO	7.62	(-0.12)	10.32	(0.07)	11.84	(-0.20)	90.19	(0.19)	89.71	(-0.29)	89.99	(-0.01)
NAPHOL	13.24	(0.06)	4.77	(-0.03)	13.25	(-0.03)	90.06	(0.06)	117.35	(0.23)	89.89	(-0.11)
PENCEN	15.92	(0.12)	6.03	(-0.03)	16.05	(0.04)	100.57	(-1.33)	111.91	(-0.69)	85.95	(0.15)
TPHBEN	7.30	(-0.31)	19.69	(-0.08)	11.27	(0.02)	90.03	(0.03)	90.00	(0.00)	89.44	(-0.56)

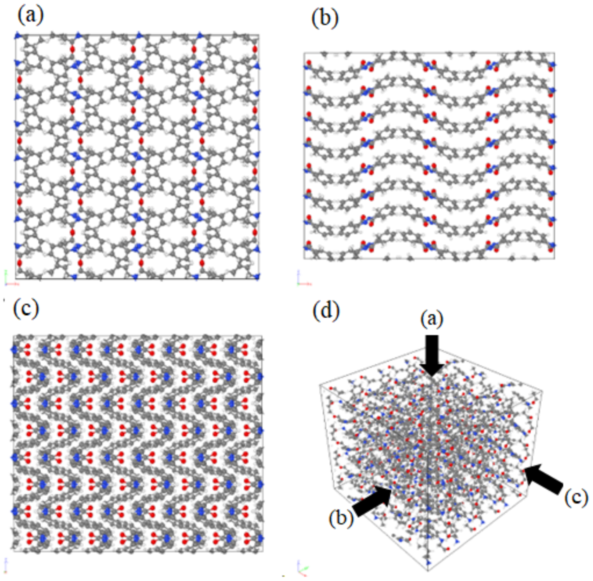


Figure 9: A supercell structure of ACANIL Crystal Structure after 1 ns of 300K MD using the SDM-optimized FF

3.4 Crystal Structure Reproducibility on Supercells and at Finite Temperature

After performing MD on supercells containing more than 3000 molecules with a time step of 1 fs, we confirmed that all crystal structures were preserved. Utilizing the SDM-optimized FFs, we performed MD at 300 K and NPT conditions, with heating from 10 K to 300 K for 0.1 ns, holding for 1 ns, and cooling to 10 K for 0.1 ns. The lattice constants of the eight molecules of crystals after simulation with supercells are summarized in Table 4. The lengths of the axes were converted to those in unit cells.

Despite the weak hydrogen bonding between the amino and carbonyl groups of ACANIL, we succeeded in reproducing the crystal structure, as depicted in Figure 9. As we employed an Amber-type FF, expressed by Eq. (1), the hydrogen bonds were not explicitly included. However, the intermolecular interactions resulting from the vdW and Coulomb forces, including the indirect correlations between the bond, bond angle, and dihedral angle, successfully reproduced the hydrogen bonds in an implicit manner. The additional seven structures obtained from the MD simulations under the same conditions are presented in Appendix 6 (A6).

4 Conclusion

Here, we propose SDM as a method for generating FFs of small organic molecules using reference data of the (i) stable monomer structure, (ii) crystal structure, (iii) lattice energy of the structure, and (iv) PESs of the dihedral angles. In principle, SDM utilized IFD to differentiate the converged structures of iterative algorithms with respect to the FF parameters, which identified the direction in which the FF parameters should be adjusted to ensure correspondence between the converged structure and the experimental crystal structure. In addition, to accurately consider the long-range interactions for optimizing the atomic charges, the Ewald method was implemented in its differentiable form.

The performance of SDM was compared with that of the FDM using eight molecules as test cases. As observed for all molecules, SDM was more accurate than FDM, which reduced the MAE of the lattice constants to 0.56 % from the corresponding error of 1.2 % for FDM. The MAE of internal atomic coordinates within the lattice was 0.16 Å for SDM and 0.22 Å for FDM. Similarly, the lattice energies were reproduced with an error of 0.14 kcal/mol for the SDM and 2.40 kcal/mol for the FDM. The MD simulations with the FFs generated by SDM for 1 ns at 300 K and NPT conditions using LAMMPS confirmed that the crystal structures were retained for all eight molecules. Thus, SDM can automate the FF optimization of small organic molecules with crystal structures as reference data. In future, studies should aim to bridge the gap between stable and experimental structures at finite temperatures.

This method is not limited to Amber-type FFs and can be easily applied to polarized FFs and special interactions explicitly described as an energy function, such as hydrogen bonding. Thus, we believe that SDM will emerge as one of the new standards for parameterizing FFs with crystal structures.

References

[1] Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I.; Brozell, S. R.; Cerutti, D. S.; Cheatham III, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E., et al. *Amber 2021*; University of California, San Francisco, 2021.

[2] Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D., et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **2022**, *271*.

[3] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I., et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry* **2010**, *31*, 671–690.

[4] Yang, Y.; Qin, J.; Liu, H.; Yao, X. Molecular dynamics simulation, free energy calculation and structure-based 3D-QSAR studies of B-RAF kinase inhibitors. *Journal of Chemical Information and Modeling* **2011**, *51*, 680–692.

[5] Wang, X.; Song, K.; Li, L.; Chen, L. Structure-Based Drug Design Strategies and Challenges. *Current Topics in Medicinal Chemistry* **2018**, *18*, 998–1006.

[6] Opo, F. A. D. M.; Rahman, M. M.; Ahammad, F.; Ahmed, I.; Bhuiyan, M. A.; Asiri, A. M. Structure based pharmacophore modeling, virtual screening, molecular docking and ADMET approaches for identification of natural anti-cancer agents targeting XIAP protein. *Scientifc Reports* **2021**, *11*, 4049.

[7] Ganguly, A.; Tsai, H.-C.; Fernández-Pendás, M.; Lee, T.-S.; Giese, T. J.; York, D. M. AMBER Drug Discovery Boost Tools: Automated Workflow for Production Free-Energy Simulation Setup and Analysis (ProFESSA). *Journal of Chemical Information and Modeling* **2022**, *62*, 6069–6083.

[8] Kippelen, B.; Brédas, J. L. Organic photovoltaics. *Energy and Environmental Science* **2009**, *2*, 251–261.

[9] Kobayashi, H.; Kobayashi, N.; Hosoi, S.; Koshitani, N. Hopping and band mobilities of pentacene, rubrene, and 2, 7-dioctyl [1] benzothieno [3, 2-b][1] benzothiophene (C8-BTBT) from first-principle calculations. *The Journal of Chemical Physics* **2013**, *014707*, 2–9.

[10] Friederich, P.; Fediai, A.; Kaiser, S.; Konrad, M.; Jung, N.; Wenzel, W. Toward Design of Novel Materials for Organic Electronics. *Advanced Materials* **2019**, *31*, e1808256.

[11] Honig, B.; Karplus, M. Implications of torsional potential of retinal isomers for visual excitation. *Nature* **1971**, *229*, 558–560.

[12] Warshel, A.; Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *Journal of the American Chemical Society* **1972**, *94*, 5612–5625.

[13] Warshel, A.; Karplus, M. Calculation of. pi.. pi.* excited state conformations and vibronic structure of retinal and related molecules. *Journal of the American Chemical Society* **1974**, *96*, 5677–5689.

[14] Karplus, M. Development of multiscale models for complex chemical systems: from H+ H2 to biomolecules (Nobel lecture). *Angewandte Chemie International Edition* **2014**, *53*, 9992–10005.

[15] Levitt, M. Birth and future of multiscale modeling for macromolecular systems (Nobel Lecture). *Angewandte Chemie International Edition* **2014**, *53*, 10006–10018.

[16] Warshel, A. Multiscale modeling of biological functions: from enzymes to molecular machines (Nobel Lecture). *Angewandte Chemie International Edition* **2014**, *53*, 10020–10031.

[17] Chung, L. W.; Sameera, W. M.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F., et al. The ONIOM Method and Its Applications. *Chemical Reviews* **2015**, *115*, 5678–5796.

[18] Zev, S.; Gupta, P. K.; Pahima, E.; Major, D. T. A Benchmark Study of Quantum Mechanics and Quantum Mechanics-Molecular Mechanics Methods for Carbocation Chemistry. *Journal of Chemical Theory and Computation* **2022**, *18*, 167–178.

[19] Dong, G.; Phung, Q. M.; Pierloot, K.; Ryde, U. Reaction Mechanism of [NiFe] Hydrogenase Studied by Computational Methods. *Inorganic Chemistry* **2018**, *57*, 15289–15298.

[20] Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: a generic force field for molecular simulations. *Journal of Physical Chemistry* **1990**, *94*, 8897–8909.

[21] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174.

[22] Qiu, Y.; Smith, D. G.; Boothroyd, S.; Jang, H.; Hahn, D. F.; Wagner, J.; Bannan, C. C.; Gokey, T.; Lim, V. T.; Stern, C. D., et al. Development and benchmarking of open force field v1. 0.0—the Parsley small-molecule force field. *Journal of Chemical Theory and Computation* **2021**, *17*, 6262–6280.

[23] Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters* **2007**, *98*, 1–4.

[24] Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer. *Nature Communications* **2021**, *12*, 1–11.

[25] Herbold, M.; Behler, J. A Hessian-based assessment of atomic forces for training machine learning interatomic potentials. *Journal of Chemical Physics* **2022**, *156*.

[26] Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chemical Science* **2017**, *8*, 3192–3203.

[27] Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nature Communications* **2019**, *10*, 1–8.

[28] Gao, X.; Ramezanghorbani, F.; Isayev, O.; Smith, J. S.; Roitberg, A. E. TorchANI: A Free and Open Source PyTorch-Based Deep Learning Implementation of the ANI Neural Network Potentials. *Journal of Chemical Information and Modeling* **2020**, *60*, 3408–3415.

[29] Devereux, C.; Smith, J. S.; Davis, K. K.; Barros, K.; Zubatyuk, R.; Isayev, O.; Roitberg, A. E. Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens. *Journal of Chemical Theory and Computation* **2020**, *16*, 4192–4202.

[30] Hao, D.; He, X.; Roitberg, A. E.; Zhang, S.; Wang, J. Development and Evaluation of Geometry Optimization Algorithms in Conjunction with ANI Potentials. *Journal of Chemical Theory and Computation* **2022**, *18*, 978–991.

[31] Schütt, K. T.; Sauceda, H. E.; Kindermans, P. J.; Tkatchenko, A.; Müller, K. R. SchNet - A deep learning architecture for molecules and materials. *Journal of Chemical Physics* **2018**, *148*.

[32] Gasteiger, J.; Groß, J.; Günnemann, S. Directional Message Passing for Molecular Graphs. International Conference on Learning Representations (ICLR). 2020.

[33] Gasteiger, J.; Giri, S.; Margraf, J. T.; Günnemann, S. Fast and Uncertainty-Aware Directional Message Passing for Non-Equilibrium Molecules. Machine Learning for Molecules Workshop, NeurIPS. 2020.

[34] Takamoto, S.; Izumi, S.; Li, J. TeaNet: Universal neural network interatomic potential inspired by iterative electronic relaxations. *Computational Materials Science* **2022**, *207*.

[35] Takamoto, S.; Shinagawa, C.; Motoki, D.; Nakago, K.; Li, W.; Kurata, I.; Watanabe, T.; Yayama, Y.; Iriguchi, H.; Asano, Y., et al. Towards universal neural network potential for material discovery applicable to arbitrary combination of 45 elements. *Nature Communications* **2022**, *13*, 2991.

[36] Thaler, S.; Zavadlav, J. Learning neural network potentials from experimental data via Differentiable Trajectory Reweighting. *Nature Communications* **2021**, *12*, 1–10.

[37] Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Physical Review B* **2019**, *99*, 1–15.

[38] Lysogorskiy, Y.; van der Oord, C.; Bochkarev, A.; Menon, S.; Rinaldi, M.; Hammerschmidt, T.; Mrovec, M.; Thompson, A.; Csányi, G.; Ortner, C., et al. Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. *npj Computational Materials* **2021**, *7*, 1–12.

[39] Bochkarev, A.; Lysogorskiy, Y.; Menon, S.; Qamar, M.; Mrovec, M.; Drautz, R. Efficient parametrization of the atomic cluster expansion. *Physical Review Materials* **2022**, *6*, 1–18.

[40] Boothroyd, S.; Madin, O. C.; Mobley, D. L.; Wang, L. P.; Chodera, J. D.; Shirts, M. R. Improving Force Field Accuracy by Training against Condensed-Phase Mixture Properties. *Journal of Chemical Theory and Computation* **2022**, *18*, 3577–3592.

[41] Boothroyd, S.; Wang, L. P.; Mobley, D. L.; Chodera, J. D.; Shirts, M. R. Open Force Field Evaluator: An Automated, Efficient, and Scalable Framework for the Estimation of Physical Properties from Molecular Simulation. *Journal of Chemical Theory and Computation* **2022**, *18*, 3566–3576.

[42] ÖzpInar, G. A.; Peukert, W.; Clark, T. An improved generalized AMBER force field (GAFF) for urea. *Journal of Molecular Modeling* **2010**, *16*, 1427–1440.

[43] Park, H.; Zhou, G.; Baek, M.; Baker, D.; DiMaio, F. Force Field Optimization Guided by Small Molecule Crystal Lattice Data Enables Consistent Sub-Angstrom Protein-Ligand Docking. *Journal of Chemical Theory and Computation* **2021**, *17*, 2000–2010.

[44] Thürlemann, M.; Böselt, L.; Riniker, S. Regularized by Physics: Graph Neural Network Parametrized Potentials for the Description of Intermolecular Interactions. *Journal of Chemical Theory and Computation* **2023**,

[45] Shalev, O.; Shtein, M. Effect of crystal density on sublimation properties of molecular organic semiconductors. *Organic Electronics* **2013**, *14*, 94–99.

[46] Marchese Robinson, R. L.; Geatches, D.; Morris, C.; Mackenzie, R.; Maloney, A. G. P.; Roberts, K. J.; Moldovan, A.; Chow, E.; Pencheva, K.; Vatvani, D. R. M. Evaluation of Force-Field Calculations of Lattice Energies on a Large Public Dataset, Assessment of Pharmaceutical Relevance, and Comparison to Density Functional Theory. *Journal of Chemical Information and Modeling* **2019**, *59*, 4778–4792.

[47] Chickos, J. S.; Gavezzotti, A. Sublimation Enthalpies of Organic Compounds: A Very Large Database with a Match to Crystal Structure Determinations and a Comparison with Lattice Energies. *Crystal Growth & Design* **2019**, *19*, 6566–6576.

[48] McDonagh, J. L.; Palmer, D. S.; van Mourik, T.; Mitchell, J. B. O. Are the Sublimation Thermodynamics of Organic Molecules Predictable? *Journal of Chemical Information and Modeling* **2016**, *56*, 2162–2179.

[49] Xie, Y.; Vandermause, J.; Sun, L.; Cepellotti, A.; Kozinsky, B. Bayesian force fields from active learning for simulation of inter-dimensional transformation of stanene. *npj Computational Materials* **2021**, *7*, 34–36.

[50] Vandermause, J.; Torrisi, S. B.; Batzner, S.; Xie, Y.; Sun, L.; Kolpak, A. M.; Kozinsky, B. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. *npj Computational Materials* **2020**, *6*, 1–11.

[51] Köfinger, J.; Hummer, G. Empirical optimization of molecular simulation force fields by Bayesian inference. *European Physical Journal B* **2021**, *94*.

[52] Liu, H.; Fu, Z.; Li, Y.; Sabri, N. F. A.; Bauchy, M. Parameterization of empirical forcefields for glassy silica using machine learning. *MRS Communications* **2019**, *9*, 593–599.

[53] Befort, B. J.; Defever, R. S.; Maginn, E. J.; Alexander, W. Machine Learning-Enabled Optimization of Force Fields for Hydrofluorocarbons. *Computer Aided Chemical Engineering* **2021**, *49*, 1249–1254.

[54] Krishnamoorthy, A.; Mishra, A.; Kamal, D.; Hong, S.; Nomura, K.; Tiwari, S.; Nakano, A.; Kalia, R.; Ramprasad, R.; Vashishta, P. EZFF: Python library for multi-objective parameterization and uncertainty quantification of interatomic forcefields for molecular dynamics. *SoftwareX* **2021**, *13*, 100663.

[55] Furman, D.; Carmeli, B.; Zeiri, Y.; Kosloff, R. Enhanced Particle Swarm Optimization Algorithm: Efficient Training of ReaxFF Reactive Force Fields. *Journal of Chemical Theory and Computation* **2018**, *14*, 3100–3112.

[56] Lombardo, T.; Hoock, J. B.; Primo, E. N.; Ngandjong, A. C.; Duquesnoy, M.; Franco, A. A. Accelerated Optimization Methods for Force-Field Parameterization in Battery Electrode Manufacturing Modeling. *Batteries and Supercaps* **2020**, *3*, 721–730.

[57] Wang, L. P.; Chen, J.; Van Voorhis, T. Systematic parameterization of polarizable force fields from quantum chemistry data. *Journal of Chemical Theory and Computation* **2013**, *9*, 452–460.

[58] Wang, L. P.; Martinez, T. J.; Pande, V. S. Building force fields: An automatic, systematic, and reproducible approach. *Journal of Physical Chemistry Letters* **2014**, *5*, 1885–1891.

[59] Qiu, Y.; Nerenberg, P. S.; Head-Gordon, T.; Wang, L. P. Systematic optimization of water models using liquid/vapor surface tension data. *Journal of Physical Chemistry B* **2019**, *123*, 7061–7073.

[60] Morado, J.; Mortenson, P. N.; Verdonk, M. L.; Ward, R. A.; Essex, J. W.; Skylaris, C. K. ParaMol: A Package for Automatic Parameterization of Molecular Mechanics Force Fields. *Journal of Chemical Information and Modeling* **2021**, *61*, 2026–2047.

[61] Baydin, A. G.; Pearlmutter, B. A.; Radul, A. A.; Siskind, J. M. Automatic Differentiation in Machine Learning: a Survey. *Journal of Machine Learning Research* **2018**, *18*, 1–43.

[62] Bradbury, J.; Frostig, R.; Hawkins, P.; Johnson, M. J.; Leary, C.; Maclaurin, D.; Necula, G.; Paszke, A.; VanderPlas, J.; Wanderman-Milne, S., et al. JAX: Composable transformations of Python+NumPy programs. 2018; `http://github.com/google/jax`.

[63] Schoenholz, S. S.; Cubuk, E. D. JAX, M.D. A framework for differentiable physics. *Advances in Neural Information Processing Systems* **2020**, *2020-Decem.*

[64] Goodrich, C. P.; King, E. M.; Schoenholz, S. S.; Cubuk, E. D.; Brenner, M. P. Designing self-assembling kinetics with differentiable statistical physics models. *Proceedings of the National Academy of Sciences of the United States of America* **2021**, *118*, 1–7.

[65] Kaymak, M. C.; Rahnamoun, A.; Hearn, K. A. O.; van Duin, A. C. T.; Merz, K. M.; Aktulga, H. M. JAX-ReaxFF : A Gradient Based Framework for Extremely Fast Optimization of Reactive Force Fields. *arXiv* **2021**, 1–29.

[66] Blondel, M.; Berthet, Q.; Cuturi, M.; Frostig, R.; Hoyer, S.; Llinares-López, F.; Pedregosa, F.; Vert, J.-P. Efficient and Modular Implicit Differentiation. **2021**, 1–29.

[67] Ewald, P. P. Ewald summation. *Annals. of Physics* **1921**, *369*, 1–2.

[68] Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[69] Steinbach, P. J.; Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry* **1994**, *15*, 667–683.

[70] JAXopt. `https://github.com/google/jaxopt` (accessed Dec 14, 2022).

[71] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H., et al. Gaussian 16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

[72] Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.

[73] Jewett, A. I.; Stelter, D.; Lambert, J.; Saladi, S. M.; Roscioni, O. M.; Ricci, M.; Autin, L.; Maritan, M.; Bashusqeh, S. M.; Keyes, T., et al. Moltemplate: A Tool for Coarse-Grained Modeling of Complex Biological Matter and Soft Condensed Matter Physics. *Journal of Molecular Biology* **2021**, *433*, 166841, Computation Resources for Molecular Biology.

[74] Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J., et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020**, *17*, 261–272.

Appendix

A1. Reproducibility of each type of energy of LAMMPS

The energy of a unit lattice obtained by LAMMPS and the proposed program are comparatively presented in Table 5. In addition to the total energy, each type of energy is compared. As observed, all types of energies were identical to each other with a maximum error of less than 1.0×10^{-3}. Relatively, the errors in FF optimization with the proposed program were subsequently smaller than those obtained with LAMMPS.

Table 5 compares the energy of the unit lattice between LAMMPS and the program in this study. In addition to the total energy, a comparison of each type of energy is also shown. All energies were identical to each other with The maximum error is less than 1.0×10^{-3} kcal/mol. The errors are sufficiently small enough to transfer the FF optimized byusing our program to LAMMPS.

Table 5: Comparison of various energies between LAMMPS and the FF generation program created in this study [kcal/mol]

program	E_{total}	E^S	$E^L - E^{\text{self}}$	E_{vdW}	E_{bond}	E_{angle}	E_{dihed}	E_{improper}
ACANIL								
LAMMPS	-846.96937	-307.06212	-310.05089	-7.69170	22.35649	64.39087	-309.48220	0.57019
This study	-846.97020	-307.06241	-310.05143	-7.69170	22.35649	64.39087	-309.48221	0.57019
difference	0.00083	0.00029	0.00054	0.00000	0.00000	0.00000	0.00001	0.00000
ANTCEN								
LAMMPS	464.24117	77.30488	-45.07281	82.78989	34.95561	56.02923	258.16100	0.07337
This study	464.24126	77.30504	-45.07289	82.78989	34.95561	56.02923	258.16100	0.07337
difference	0.00009	0.00016	0.00007	0.00000	0.00000	0.00000	0.00000	0.00000
BENZAC								
LAMMPS	30.13130	60.25193	-165.00981	102.52778	57.52837	5.30274	-30.47399	0.00428
This study	30.13149	60.25222	-165.00991	102.52779	57.52837	5.30274	-30.47399	0.00428
difference	0.00019	0.00028	0.00010	0.00000	0.00000	0.00000	0.00000	0.00000
BIPHEN								
LAMMPS	122.75928	5.09149	-2.89838	71.24001	26.75678	13.56645	9.00133	0.00160
This study	122.75928	5.09149	-2.89839	71.24001	26.75678	13.56645	9.00133	0.00160
difference	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
BPHENO								
LAMMPS	-1230.34900	40.78411	-29.84355	29.88174	16.31054	4.63192	-1292.11520	0.00146
This study	-1230.34895	40.78419	-29.84358	29.88174	16.31054	4.63192	-1292.11521	0.00146
difference	0.00005	0.00008	0.00003	0.00000	0.00000	0.00000	0.00001	0.00000
NAPHOL								
LAMMPS	31.66049	-6.22532	-58.50986	24.88325	16.69255	78.60605	-23.94086	0.15467
This study	31.66041	-6.22530	-58.50996	24.88325	16.69255	78.60605	-23.94086	0.15467
difference	0.00008	0.00002	0.00010	0.00000	0.00000	0.00000	0.00000	0.00000
PENCEN								
LAMMPS	-577.03574	107.29435	-80.22260	134.51915	41.65793	236.04115	-1016.32620	0.00052
This study	-577.03585	107.29463	-80.22299	134.51916	41.65793	236.04115	-1016.32624	0.00052
difference	0.00011	0.00028	0.00039	0.00001	0.00000	0.00000	0.00004	0.00000
TPHBEN								
LAMMPS	94.82632	29.68179	-18.89986	87.29577	33.82538	21.35936	-58.44956	0.01345
This study	94.82637	29.68185	-18.89987	87.29577	33.82538	21.35936	-58.44956	0.01345
difference	0.00005	0.00006	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000

A2. Weights multiplied to each element in total evaluation function

(1) w^{M} was set to $1.0/N_{\mathrm{atom}}$, where N_{atom} denotes the number of atoms in the single molecule.

(2) w^{C} was defined as $1.0/N_{\mathrm{atom}}/N_{\mathrm{mol}}$, where N_{mol} indicated the number of molecules in a unit cell. The denominator N_{mol} is introduced to maintain the balance of multiple evaluation function types L^{M} and L^{C}.

(3) w^{E} was set $0.25/N_{\mathrm{atom}}$ (kcal/mol)$^{-2}$. Unlike structure matching with L^{M} and L^{C}, L^{E} represents an energy-matching function in a distinct unit of L^{M} and L^{C}. The weights permit a uniform treatment of the evaluation functions. For instance, a bond with the typical force constant of $K_r = 200$ (kcal/mol/Å^2). the difference of the length of 0.1 Å (=the 1 unit after regularization) increases $w^{\mathrm{M}}L^{\mathrm{M}}$ by $1.0/N_{\mathrm{atom}}$. This difference of 0.1 Å affects 4.0 (kcal/mol)^2 ($= 1.0/N_{\mathrm{atom}}[200 \times 0.1^2]$) to L^{E}, or $1.0/N_{\mathrm{atom}}$ to $w^{\mathrm{E}}L^{\mathrm{E}}$. This is identical to the variations in $w^{\mathrm{M}}L^{\mathrm{M}}$.

(4) $w^{\mathrm{M,FDM}}$ was set to $1.0/N_{\mathrm{atom}}/20^2$ (kcal/mol/Å)$^{-2}$. The deviation of the bond length of 0.1Å from the equilibrium length with the typical force constant of $K_r = 200$ (kcal/mol/Å^2) generates a residual force of 20 (kcal/mol/Å) $[= 1/2 \times 2 \times K_r \times 0.1]$ for each atom on the both sides of the bond. The difference of 0.1 Å affects 400 (kcal/mol)^2 Å^{-2} to $L^{\mathrm{M,FDM}}$, or $1.0/N_{\mathrm{atom}}$ to $w^{\mathrm{M,FDM}}L^{\mathrm{M,FDM}}$.

(5) Similar to $w^{\mathrm{M,FDM}}$, $w^{\mathrm{C,FDM}}$ was set to $1.0/N_{\mathrm{atom}}/N_{\mathrm{mol}}/20^2$ (kcal/mol/Å)$^{-2}$.

A3. Boundary of each FF parameter type

Both FDM and SDM methods employ bounds to limit the range of the FF parameters as shown in Table 6. They are primarily used to avoid the structure optimization errors caused by excessively large or negligibly small FF parameter values, especially in the initial stages of FF parameter optimization.

Table 6: Inequality constaraints of FF optimization for maintaining the parameters within the physically valid values

Type	Inequality conditions
Bond force constant K_r	$0 < K_r < 2K_{r0}$
Equilibrium bond length r_{eq}	$0 < r_{\mathrm{eq}} < 2r_{\mathrm{eq}0}$
Equilibrium angle θ_{eq}	$-180° < \theta_{\mathrm{eq}} < +180°$
dihedral force constant V_n	$-5\mathrm{kcal/mol} < V_n < +5\mathrm{kcal/mol}$
vdW well depth ϵ	$0\mathrm{kcal/mol} < \epsilon < 2\mathrm{kcal/mol}$
vdW length σ	$0 < \sigma < 2\sigma_0$
Charge q	$q_0 - 0.5e < q < q_0 + 0.5e$

A4. All PESs calculated with SDM-optimized FF

As the evaluation function of PESs (Eq. (17)) does not include structure optimizations and compares only the energy of specific and fixed structures, we focus on the validity of PESs derived from SDM instead of a comparison between FDM and SDM. The reproducibility of PES ϕ_A is indicated in the inset of Figure 10(a). The PES of GAFF is unsuitable for MD simulations owing to the discrepancy between the most stable point and the reference data. Conversely, SDM reproduces the agreement of the minima as well as the curvature near the minima and the height of the peak near from 90°. As plotted in Figure 10(b), the vdW and Coulomb potentials contribute to these reproductions. In addition to the bond between the aromatic carbon atom and nitrogen atom, the intramolecular hydrogen bond between the oxygen atom of the carbonyl group and the hydrogen atom at the *ortho*-position of the benzene ring must be considered for the dihedral angle. In SDM, these effects are effectively represented by a combination of the FF parameters. Similar to ACANIL, SDM could accurately reproduce the minima and its curvature for all PES in the eight molecules. The results of the optimized PESs are displayed in the following figures, *i.e.*, Figures 10–16.

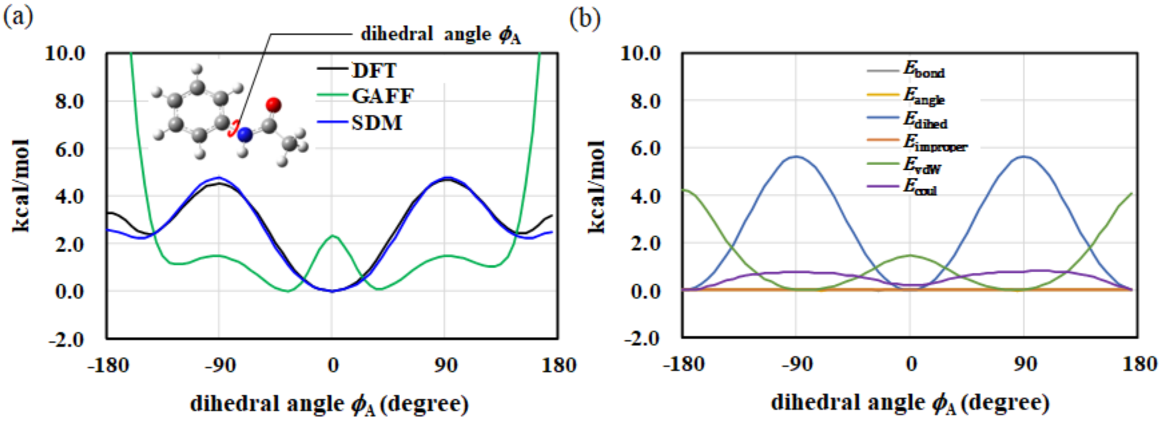


Figure 10: PES of dihedral angle ϕ_A in ACANIL (a) comparison of reference data, GAFF, and SDM, (b) each energy type in FF derived using SDM

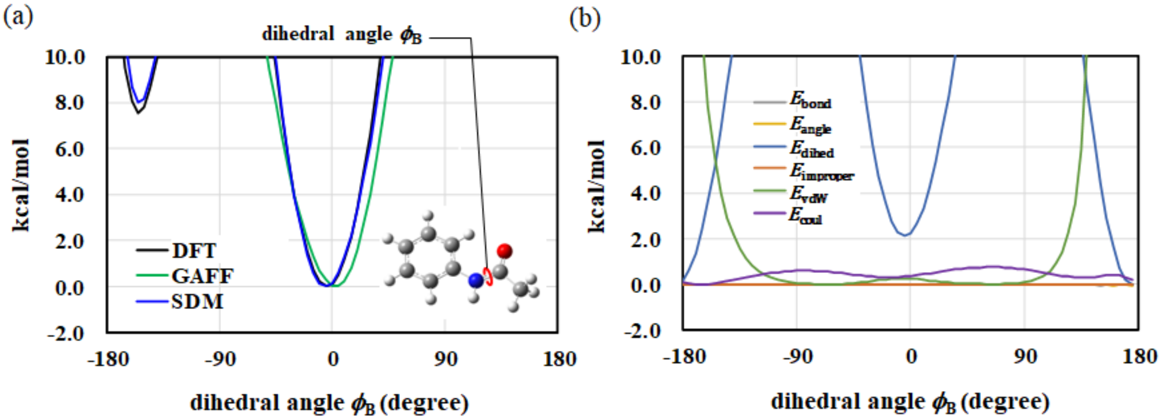


Figure 11: PES of dihedral angle ϕ_B in ACANIL (a) comparison of reference data, GAFF, and SDM, (b) each energy type in FF derived using SDM

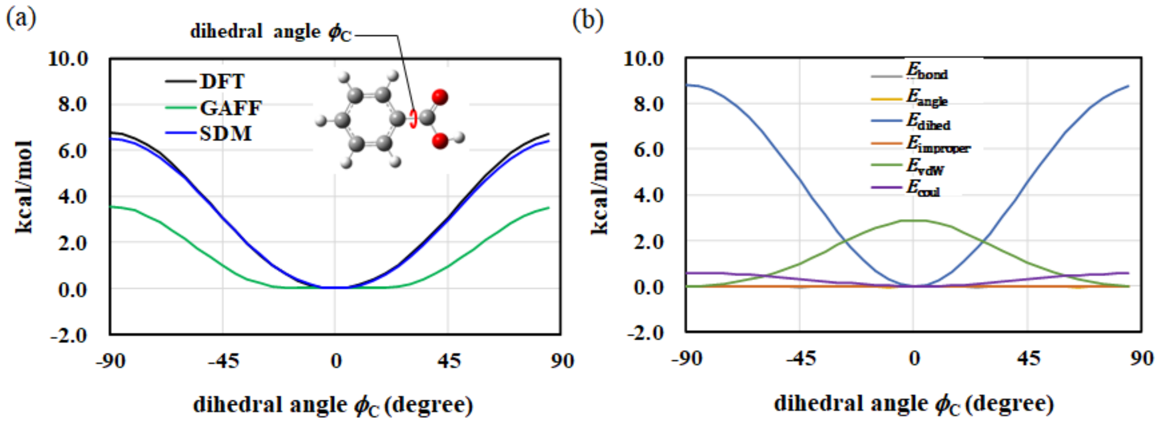


Figure 12: PES of dihedral angle ϕ_C in BENZAC (a) comparison of reference data, GAFF, and SDM, (b) each energy type in FF derived using SDM

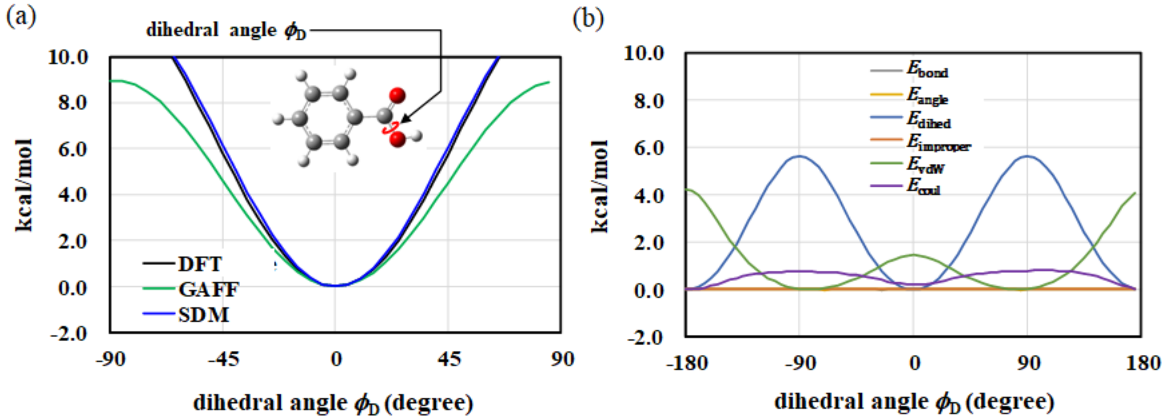


Figure 13: PES of dihedral angle ϕ_D in BENZAC (a) comparison of reference data, GAFF, and SDM, (b) each energy type in FF derived using SDM

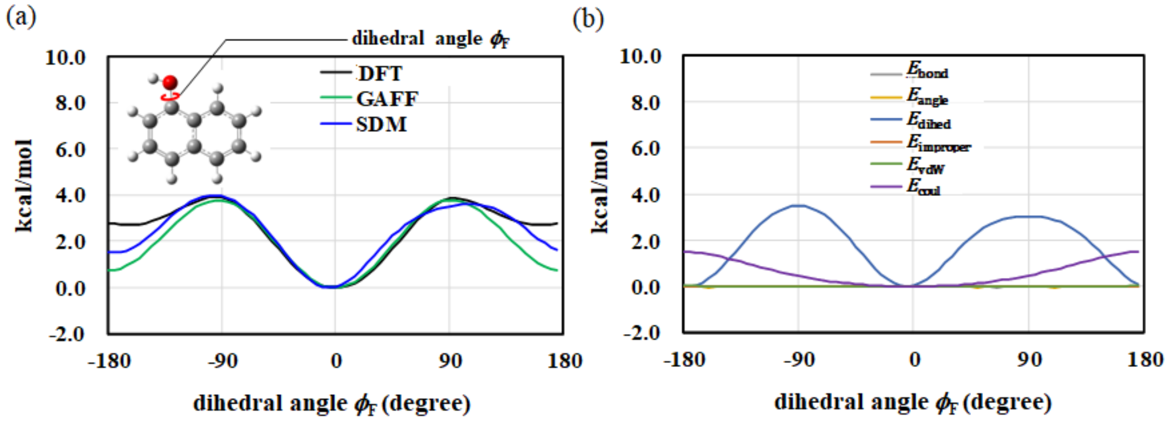


Figure 14: PES of dihedral angle ϕ_E in BPHENO (a) comparison of reference data, GAFF, and SDM, (b) each energy type in FF derived using SDM

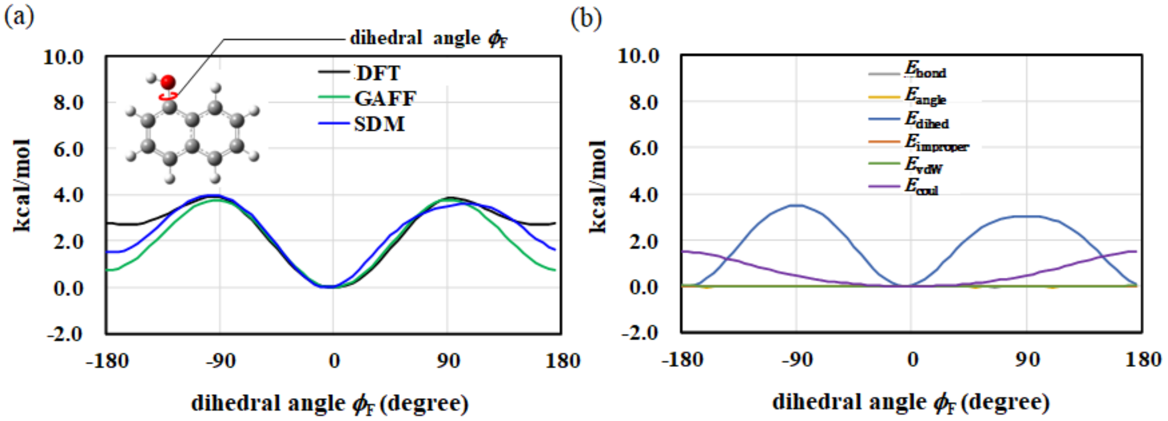


Figure 15: PES of dihedral angle ϕ_F in NAPHOL (a) comparison of reference data, GAFF, and SDM, (b) each energy type in FF derived using SDM

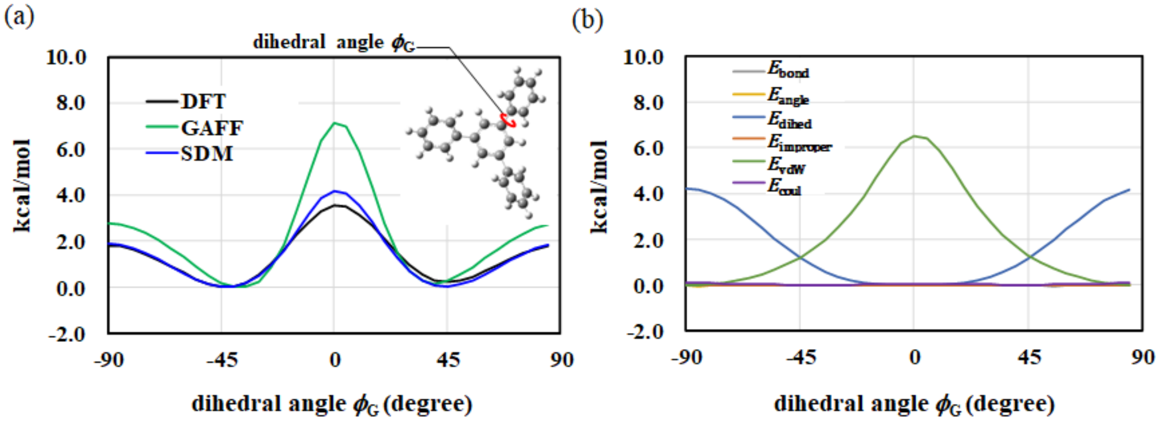
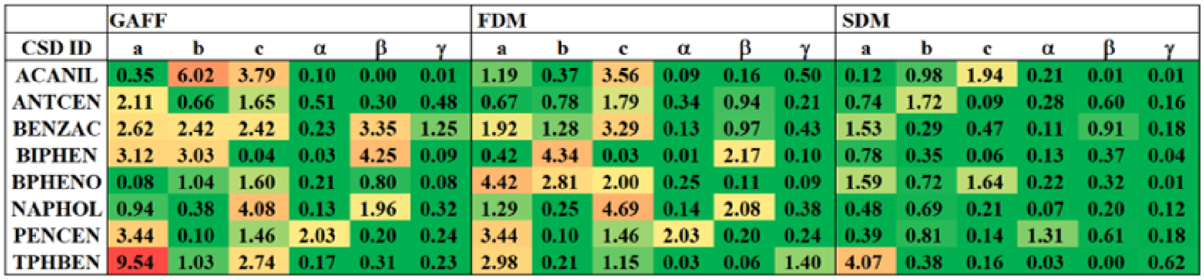


Figure 16: PES of dihedral angle ϕ_G in TPHBEN (a) comparison of reference data, GAFF, and SDM, (b) each energy type in FF derived using SDM

A5. Heatmaps of errors in lattice constants

The heatmap of the lattice constants errors optimized with the initial GAFF and the FFs optimized by FDM and SDM are illustrated in Figure 17. The background color is assigned according to each value, where green indicates a smaller evaluation function value and red denotes a larger value.



	GAFF						FDM						SDM					
CSD ID	a	b	c	α	β	γ	a	b	c	α	β	γ	a	b	c	α	β	γ
ACANIL	0.35	6.02	3.79	0.10	0.00	0.01	1.19	0.37	3.56	0.09	0.16	0.50	0.12	0.98	1.94	0.21	0.01	0.01
ANTCEN	2.11	0.66	1.65	0.51	0.30	0.48	0.67	0.78	1.79	0.34	0.94	0.21	0.74	1.72	0.09	0.28	0.60	0.16
BENZAC	2.62	2.42	2.42	0.23	3.35	1.25	1.92	1.28	3.29	0.13	0.97	0.43	1.53	0.29	0.47	0.11	0.91	0.18
BIPHEN	3.12	3.03	0.04	0.03	4.25	0.09	0.42	4.34	0.03	0.01	2.17	0.10	0.78	0.35	0.06	0.13	0.37	0.04
BPHENO	0.08	1.04	1.60	0.21	0.80	0.08	4.42	2.81	2.00	0.25	0.11	0.09	1.59	0.72	1.64	0.22	0.32	0.01
NAPHOL	0.94	0.38	4.08	0.13	1.96	0.32	1.29	0.25	4.69	0.14	2.08	0.38	0.48	0.69	0.21	0.07	0.20	0.12
PENCEN	3.44	0.10	1.46	2.03	0.20	0.24	3.44	0.10	1.46	2.03	0.20	0.24	0.39	0.81	0.14	1.31	0.61	0.18
TPHBEN	9.54	1.03	2.74	0.17	0.31	0.23	2.98	0.21	1.15	0.03	0.06	1.40	4.07	0.38	0.16	0.03	0.00	0.62

Figure 17: Heatmap of lattice constants errors optimized with GAFF, and the FF optimized with FDM, and SDM (Errors (%) in axis lengths a,b, and c; and angles between axis of b–c, c–a, and a–b: $\alpha, \beta, and \gamma$)

A6. All Crystal Structures Reproducibility on supercells and Finite Temperature

The following Figure 18–24 include all supercell structures after the MD with the same condition described in the section 3.4. The figure for ACANIL has already shown in Figure 9.

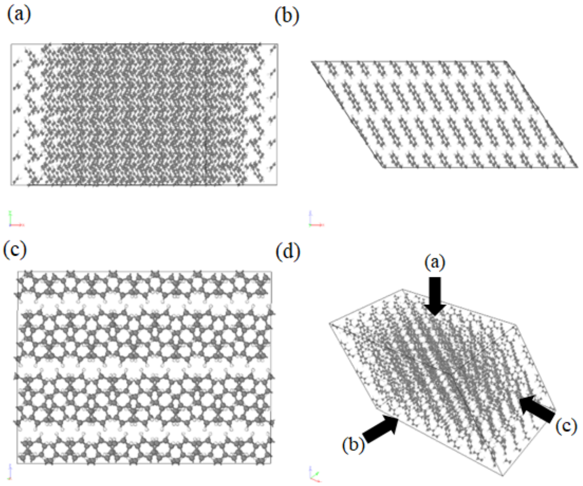


Figure 18: A supercell structure of ANTCEN after 1 ns of MD at 300 K with SDM-optimized FF

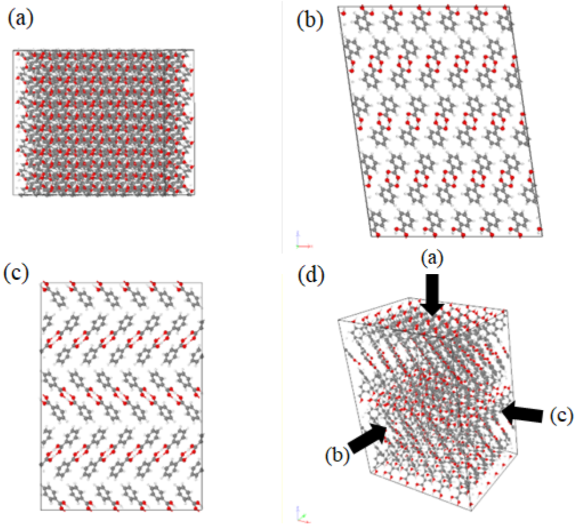


Figure 19: A supercell structure of BENZAC after 1 ns of MD at 300 K with SDM-optimized FF

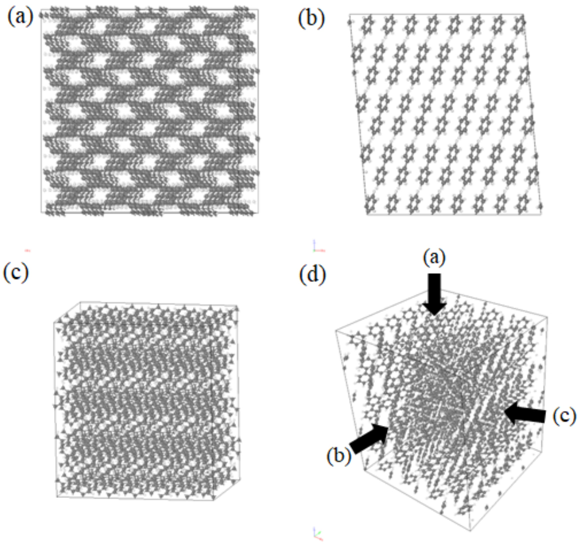


Figure 20: A supercell structure of BIPHEN after 1 ns of MD at 300 K with SDM-optimized FF

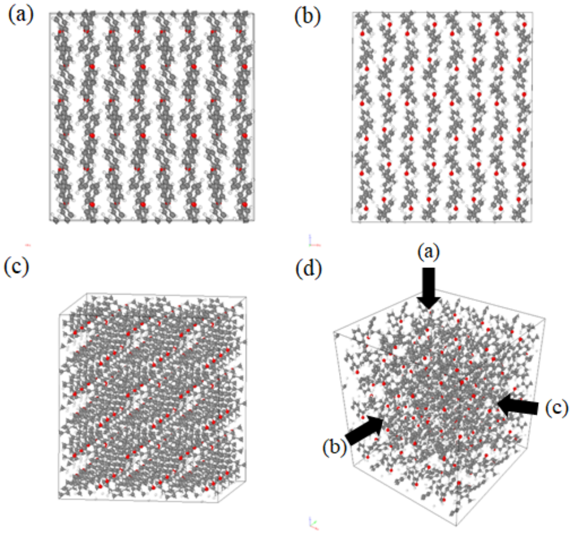


Figure 21: A supercell structure of BPHENO after 1 ns of MD at 300 K with SDM-optimized FF

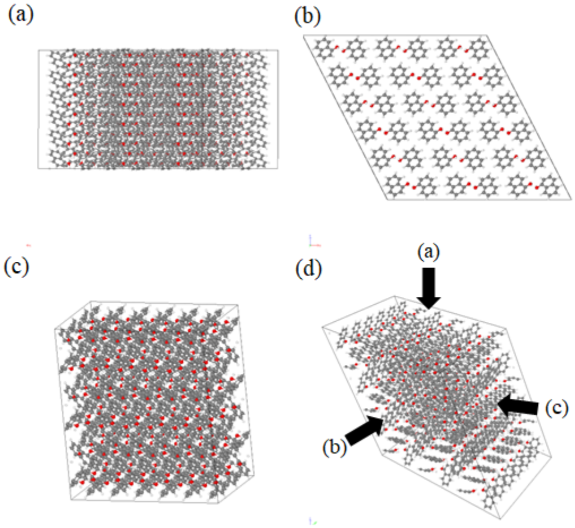


Figure 22: A supercell structure of NAPHOL after 1 ns of MD at 300 K with SDM-optimized FF

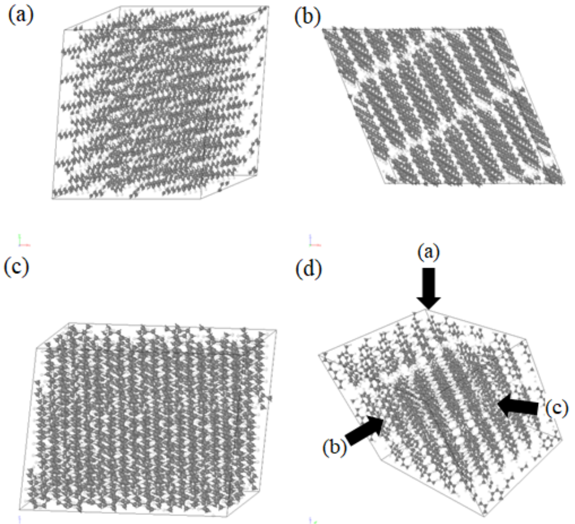


Figure 23: A supercell structure of PENCEN after 1 ns of MD at 300 K with SDM-optimized FF

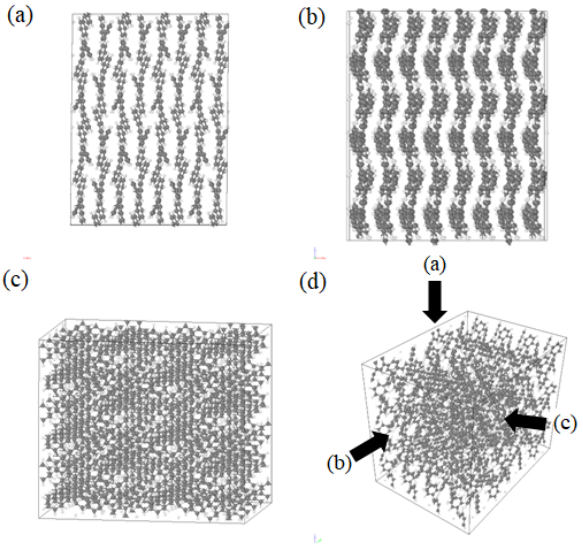


Figure 24: A supercell structure of TPHBEN after 1 ns of MD at 300 K with SDM-optimized FF