

Polymer Structure Prediction from First Principles

Tran Doan Huan* and Rampi Ramprasad*

Cite This: *J. Phys. Chem. Lett.* 2020, 11, 5823–5829

Read Online

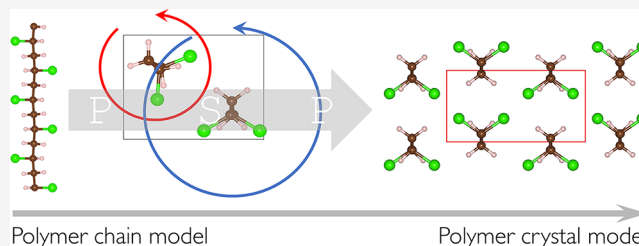
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Developing a large database of polymers structures and properties, for which suitable polymer structural models are a prerequisite, is critical for polymer informatics. We present a simple strategy, referred to as polymer structure predictor (PSP), for predicting the crystal structural models of linear polymers, given their chain-level atomic connectivity information. The PSP, which was designed specifically for polymers, relies on properly defining and sampling the configuration space. Using this approach, we have successfully recovered eight known crystal structures of six common linear polymers, including polyethylene, polyvinylidene fluoride, poly(vinyl chloride), poly(*p*-phenylene sulfide), polyacrylonitrile, and poly-2,5-benzoxazole, while discovering some new stable structures of three of them, i.e., polyvinylidene fluoride, polyacrylonitrile, and poly(*p*-phenylene sulfide). The PSP is very simple, highly scalable, suitable for automatic workflows, and comparable to the best major structure prediction method in terms of efficiency in polymer crystal structure prediction. Although challenges in comprehensively accounting for possible chain-level conformations remain, the PSP will be very useful in efficiently generating polymer data and strengthening the emerging polymer informatics ecosystems.



Understanding the atomic structure of solid materials is an important topic in materials research,¹ for which experimental techniques such as X-ray diffraction are the traditional approaches. In the last two decades, computational crystal structure prediction methods have emerged,² playing an important role in materials discovery and design. Given a chemical composition, low-energy (and, possibly, the most stable) atomic structures are predicted and the associated properties are computed, identifying viable hypothetical materials for synthesizing and testing.^{3–9} Material structure prediction is an optimization problem and involves the search for the lowest energy arrangements of a given number of atoms in a periodically repeating unit cell, whose volume and shape can change. Major structure prediction methods include the Universal Structure Predictor: Evolutionary Xtallography (USPEX),¹⁰ the minima-hopping method (MHM),^{11,12} CALYPSO,¹³ and the *ab initio* random structure search (AIRSS) method,¹⁴ using which numerous new materials have been predicted computationally and then confirmed experimentally.¹⁵

Polymer structures are highly complex, typically containing crystalline, semicrystalline, and amorphous domains existing simultaneously in which the degree of crystallinity could range from 0 up to $\approx 80\%$.¹⁶ At the first-principles level of computations, modeling polymers as crystals is widely accepted as the most reasonable approach because many important electronic and electrical properties, e.g., band gap, dielectric response, and stability, can be approximated reasonably.^{3,9,17} However, typical polymer crystal models contain hundreds of atoms in a unit cell, and if such a staggering space is explored

directly by the aforementioned (unconstrained) methods, the probability of landing at the desired polymer is very low. Despite some efforts to adapt these methods,^{4–7,9} predicting polymer structure remains enormously more challenging than materials of other classes, e.g., hard materials. We envision that the best approach is to start from the polymer atomic connectivity, generally known in terms of simplified molecular-input line-entry system (SMILES),¹⁸ in order to progressively create models of increasing sophistication, e.g., monomers, polymer chains, crystals, and perhaps amorphous models. Within this multistep pathway, which we will refer to as the polymer structure predictor (PSP) (visualized in Figure 1), the atomic connectivity is continuously monitored and strictly preserved, thus reducing the search space dimensionality and improving the efficiency of the method.

This paper focuses on a critical step of the envisioned PSP pathway and presents a simple procedure for computationally predicting linear polymer crystal structures by packing 1-dimensional polymer chains into low-energy 3-dimensional structures. Within this approach, also referred to as the PSP from now on for the sake of simplicity, possible degrees of freedom of the problem are properly defined and scanned, creating a pool of initial crystal structures, which are then optimized using density

Received: May 19, 2020

Accepted: July 1, 2020

Published: July 1, 2020

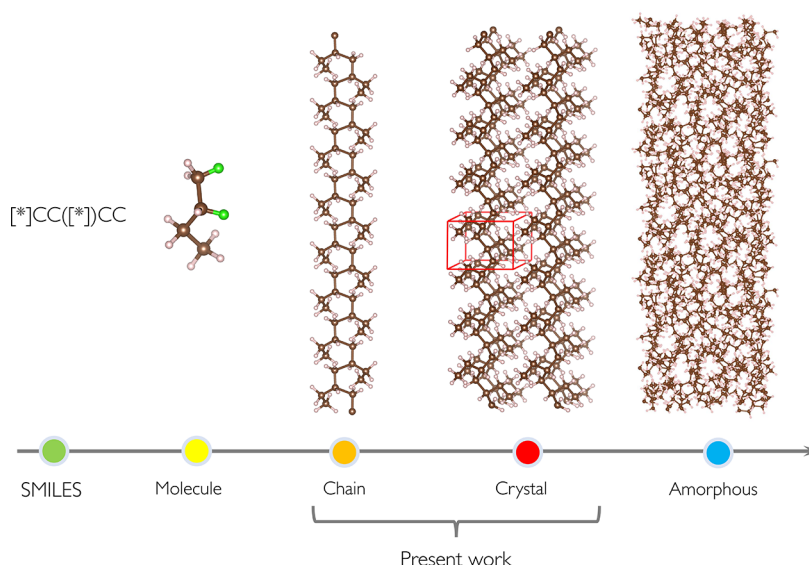


Figure 1. Visualization of the proposed polymer structure predictor. Starting from polymer SMILES, models of increasing sophistication, e.g., molecules, polymer chains, crystals, and amorphous models, will be progressively created. In the molecular model, green spheres indicate the positions at which adjacent monomers are linked while, in the crystal model, the red box is the unit cell of the crystal.

functional theory (DFT)^{19,20} computations for identifying the lowest energy candidates. We demonstrate this approach by predicting the structures of six linear common (benchmark) polymers, including polyethylene (PE),²¹ polyvinylidene fluoride (PVDF),^{22,23} poly(vinyl chloride) (PVC),²⁴ poly(*p*-phenylene sulfide) (PPS),²⁵ polyacrylonitrile (PAN),²⁶ and poly-2,5-benzoxazole (ABPBO).²⁷ While all eight targeted (experimentally known) structures of these polymers were recovered, we also predict some new structures of β -PVDF, PAN, and PPS, which are lower than the known phases in energy. Although the PSP relies on a good starting point, i.e., a reasonable polymer chain conformation, it requires little human intervention and is extremely efficient and scalable. Therefore, it is suitable for autonomously exploring the polymer space,^{5–8} generating polymer data,⁹ and promoting the emerging polymer informatic ecosystems.^{28,29}

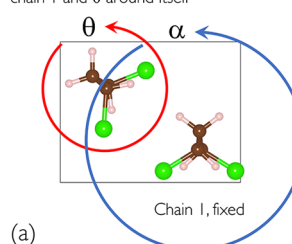
The proposed PSP strategy, which, in the present contribution, aims at constructing crystal models of a linear polymer from its chain model, involves two steps. First, the configurational space is defined and sampled, and then, the samples (the initial crystal structures) are optimized for determining the lowest energy candidates. Compared to AIRSS,¹⁴ which also relies on sampling the configurational space, this approach involves much fewer and more independent degrees of freedom; thus, the sampling step is significantly more efficient.

The polymer chains, introduced in Figure 1 and used as the input of this strategy, are infinite along the *z* axis (due to the periodicity) and isolated along the *x* and *y* directions (ensured by a large enough vacuum region). This model is also reasonably good in capturing some electronic properties, e.g., band gap,³⁰ of linear polymers because the interchain interactions are generally very weak, hardly contributing to the electronic structure. Moreover, polymer chains can be created in a high-throughput manner, given that some technical challenges of this process can be resolved appropriately. Therefore, polymer chains were extensively used in recent polymer design works,^{3,5} especially those that involve electronic band gap estimation. For evaluating truly long-length- (preferably macroscopic-scale) scale proper-

ties of polymers, crystal (and, if possible, amorphous) polymer models are required.

Polymer crystal models typically consist of at least two polymer chains packed in a low-energy (thermodynamically stable) manner because the chain dipole moment in the *xy* plane is generally nonzero. Therefore, in the structure prediction procedure described here, we consider two polymer chains packed in a variety of ways, creating a pool of initial (periodic) crystal structures. The central idea is that for two parallel polymer chains, the position of the second chain could be specified with respect to the first chain using three degrees of freedom, as visualized in Figure 2 and described below. We note

Chain 2, rotated by α around chain 1 and θ around itself



Chain 2, shifted by *d* wrt chain 1

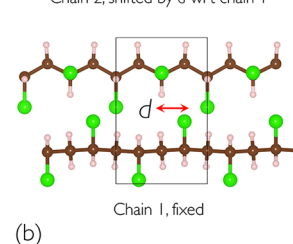


Figure 2. Two polymer chains are packed together in a simulation cell, shown as black boxes. By fixing chain 1, rotating chain 2 by α around chain 1, shifting chain 2 by *d* along chain 1, and rotating chain 2 by θ around itself, we constructed a pool of *initial* crystal structures. PVC is used in this figure, where carbon, hydrogen, and chloride atoms are shown in brown, pink, and green.

that if the dipole moment points along the *z* direction, the antiparallel orientation of the second chain with respect to the first chain should also be considered.

Assuming that the first chain is fixed and oriented along the *z* axis, the second chain is created by (1) making a copy of the first chain and placing it parallel to the first chain at a large enough distance from its origin, (2) rotating the second chain by an angle of α around the first chain, (3) rotating it by an angle of θ around itself, (4) translating it by a distance of *d* along the *z* axis

with respect to the first chain, (5) adjusting the interchain distance (by rigidly shifting the second chain) so that no interatomic distance is smaller than a cutoff value, chosen to be 2 Å in this work, and (6) finally, adjusting the periodical cell dimensions along the x and y directions so that the smallest interatomic distance between the obtained structure (with two polymer chains) and its periodic images is approximately but no less than 2 Å. This cutoff distance is large enough for not creating any unwanted interchain bonds but also small enough so that the subsequent first-principles optimizations are efficient.

Among the three degrees of freedom, α and θ range from 0 to 2π while d can take values between 0 and c , the repeat unit cell length along the z axis. Therefore, by choosing a variety of accepted α , θ , and d values, the configurational space may be adequately sampled. In particular, if we select 6 values for each degree of freedom, the number of obtained structures to be optimized is $6 \times 6 \times 6 = 216$, being generally much less than the number of DFT optimizations required for a typical crystal structure search. Moreover, unlike other structure search methods, the present approach preserves the known atomic connectivity throughout the process, and is specifically geared for polymers.

We demonstrate the efficiency of the proposed PSP strategy by performing searches for low-energy structures of six linear polymers, i.e., PE, PVDF, PVC, PPS, PAN, and ABPBO.²⁷ These simple polymers were selected from a limited set of just a few dozen polymers whose crystal structures have been resolved experimentally.⁹ In spite of their simplicity, predicting their crystal structure is, as will be shown in the next part of this paper, very challenging for established structure prediction methods; thus, they are suitable for this demonstration. The known crystal structures of these polymers, which are summarized in Table 1,

Table 1. Summary of Seven Polymers Crystal Structure Searches, Which Recovered All the Targeted Known Phases and Discovered Some New (Energetically More Stable) Phases of PE, PVDF, PVC, PPS, PAN, and ABPBO

search	polymer	known phase	new phase?
1	PE	orthorhombic ²¹ monoclinic ⁴⁰	no
2	PVDF	orthorhombic (β phase) ^{22,23}	yes
3	PVDF	monoclinic (δ phase) ²²	no
4	PVC	orthorhombic ²⁴	no
5	PPS	orthorhombic ²⁵	yes
6	PAN	orthorhombic ²⁶	yes
7	ABPBO	orthorhombic ²⁷	no

are the targets of these searches. Starting from their (known) correct atomic chain structures, 216 initial structures were generated for each polymer (6 equispaced values were selected for each of α , θ , and d). In the case of PVDF, the significantly different atomic chain models of its β and δ phases were used; thus, the total number of structure searches using the PSP is seven.

Given the initial atomic structures obtained from the sampling procedure, first-principles calculations were used to optimize them. In the work reported in this paper, such calculations were performed at the level of DFT using the projector augmented wave method³¹ as implemented in Vienna *Ab initio* Simulation Package (VASP).^{32–35} A basis set of plane waves with kinetic energy up to 400 eV was used to represent the Kohn–Sham orbitals while the Brillouin zones were sampled by Monkhorst–

Pack k -point meshes that are equally spaced by 0.2 \AA^{-1} and the number of k -points along any axis is not less than 3. The van der Waals (vdW) dispersion interactions, which were known³⁶ to be critical for describing the nonbonding interchain interactions of polymers, were estimated with the nonlocal density functional vdW-DF2.³⁷ Refitted Perdew–Wang 86,³⁸ the GGA XC functional associated with vdW-DF2, was used for the geometry optimization, for which convergence was assumed when the atomistic forces become less than 0.01 eV/\AA . This is the procedure previously adopted in our polymer data development work.⁹

During the optimization process, both the atomic and cell degrees of freedom are slowly relaxed by setting POTIM = 0.1, thus significantly lowering the optimization step widths. Generally, the initial interchain distance is large, and thus, the polymer chains are brought together. If the input polymer chain includes unsaturated bonds, extra bonds may be developed, linking different chains when they come closer during the optimization. These interchain bonds are unwanted because the resulting material is not the linear polymer we target at. In order to identify such the situation, we computed the atomic (bond-based) fingerprints¹⁷ of the crystal structures during the optimization process and compared the results with that of the initial polymer chain and/or of the original SMILES of the polymer.²⁸ Such fingerprints are sensitive with the interchain bond formation, and thus, this particular entry will be flagged and inspected visually. The results of the searches are summarized in Figures 3 and 4 and in the Supporting Information.

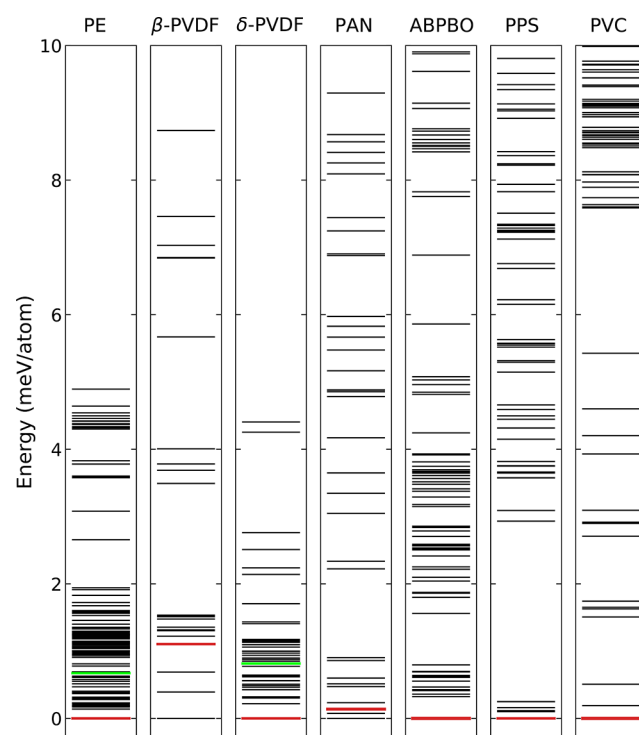


Figure 3. Low-energy spectra of the predicted polymer structures. The experimentally known structures, summarized in Table 1 and recovered in this search, are shown in colors. For PE, red and green lines represent the orthorhombic and monoclinic phases, respectively, while for the δ -PVDF search, the red and green lines represent the δ and α phase of PVDF.

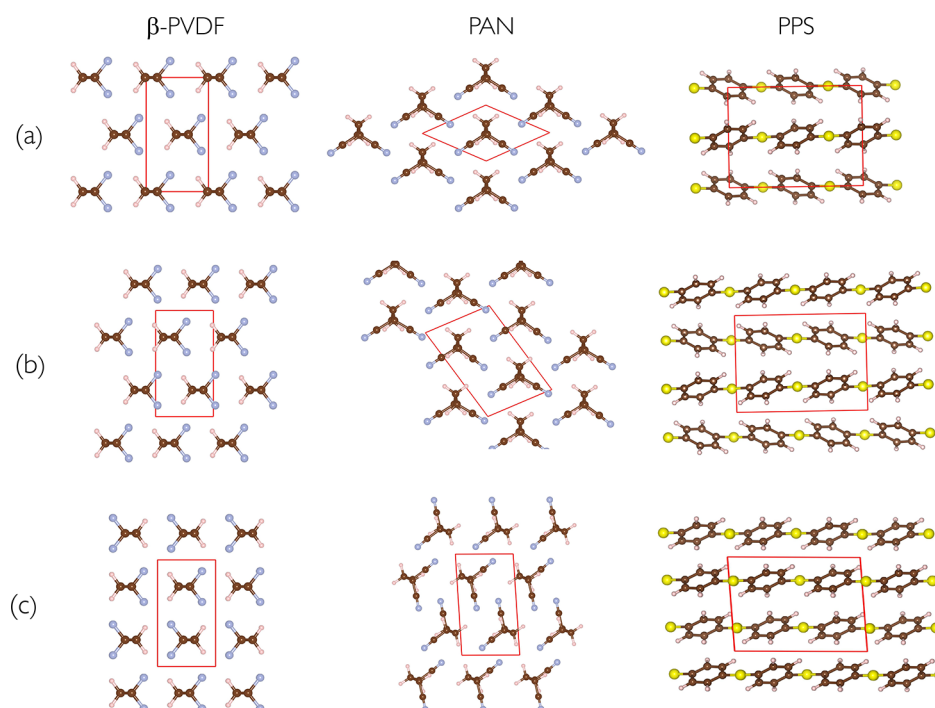


Figure 4. Low-energy structure searches for β -PVDF, PAN, and PPS recovered the experimentally known phases (a) of these polymers (the predicted structures are shown in (b)) and discovered new structures at lower energies (c). The unit cells are shown in red while carbon, hydrogen, fluoride, and sulfur atoms are given in brown, pink, cyan, and yellow, respectively.

Polyethylene (PE), a model polymer for numerous computational studies,^{41–44} has at least two known phases, i.e., the orthorhombic ground state structure²¹ and a monoclinic high-pressure structure.^{39,40} These phases consist of the same zigzag chains but they are different in the chain-packing pattern. Our PSP structure search recovered both of them as well as many other phases, as shown in Figure 3. These energetically competing structures are different from the orthorhombic and monoclinic phases in terms of the orientation and the translated distance of the second chain with respect to the first chain. Similarly, among at least four phases of PVDF, the nonplanar chain conformations of the paraelectric α phase and the ferroelectric δ phase are very similar, while the zigzag planar chain conformation of the β phase is significantly different (see Figure 5 for a visualization).^{22,45} As a result, the search starting

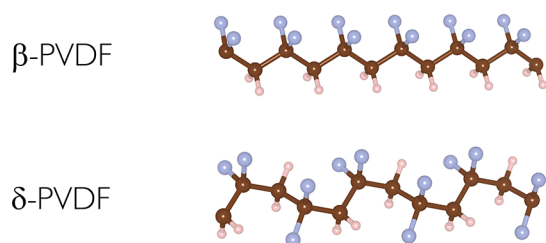


Figure 5. Polymer chains of β -PVDF and δ -PVDF. Carbon, hydrogen, and fluoride atoms are given in brown, pink, and cyan, respectively.

from the δ -PVDF chain recovered both the α and δ phases of PVDF but not the β phase. However, the search for β -PVDF recovered this phase, but not the α and δ phases. In the other searches for PAN, ABPBO, PPS, and PVC, all the targeted (known) crystal structures of these polymers were recovered as expected while a large number of energetically competing structures were also revealed. Detailed comparison between the

known and recovered structures, given in terms of structure visualization and atomic distances, can be found in the Supporting Information.

Remarkably, the new crystal structures discovered for β -PVDF, PAN, and PPS are lower in energy than the known phases by ≈ 1.0 , 0.1, and 0.01 meV/atom, respectively. These small energy differences are about 1–2 orders smaller than the typical energy differences between low-lying structures of hard materials^{46,47} and are comparable with the numerical uncertainty of the DFT calculations performed. Nevertheless, the packing patterns of the new structures discovered from the polymer chains of β -PVDF, PAN, and PPS are significantly different from the known structures, as visualized in Figure 4 and confirmed by the simulated X-ray diffraction patterns, which are also available in the Supporting Information.

We further evaluated the efficiency of the presented approach by performing a comparative study involving three major structure prediction methods, i.e., AIRSS,¹⁴ USPEX,¹⁰ and MHM.^{11,12} In AIRSS, a large number of initial structures were generated with minimal constraints and then optimized. USPEX employs the evolution algorithm to evolve an initial population of structures over successive generations, preserving the best “genomes”. When USPEX is used for polymers, their structures are generated by joining predefined rigid “building blocks”, hereby enforcing the required atomic-level connectivity.⁴ In such a constrained global optimization method, the polymer building blocks are generally small; thus, search space dimensionality is still much larger than that of the PSP. Finally, MHM, which is an essentially unconstrained search strategy, provides a mechanism for overcoming the energy barriers during the exploration of the energy landscape. Structure predictions using MHM are typically performed in several parallel searches, each of them starts from a different initial structure.

For definiteness, we set the objective of this assessment to determine whether or not AIRSS, USPEX, and MHM can correctly recover both the orthorhombic and monoclinic phases of PE. And, in case they do, we would like to determine how many DFT optimizations (using VASP) are needed. Using the AIRSS package,¹⁴ we generated 500 initial structures, requiring interatomic distances to be greater than 1.3 Å, and then optimized them. Between two searches using USPEX, USPEX-1 assumed no prior knowledge of atomic connectivity while USPEX-2 adopted the strategy developed in ref 4 for polymers, considering the (resolved) atomic structure of $-\text{CH}_2-\text{CH}_2-$ as the building block. USPEX-1 and USPEX-2 are terminated if there is no improvement during the last 20 generations (population size of a generation is 20) or if the maximum number of 50 generations is reached. Similarly, the first search (MHM-1) using MHM was started from a randomly generated arrangement of four C atoms and eight H atoms in a unit cell while the second search (MHM-2) involves 10 parallel (sub)searches, starting from 10 initial structures selected from 216 candidate structures generated in this work. Two of them are the “right” candidates (i.e., they have led to the orthorhombic and monoclinic phases of PE in the PSP search), while the others were randomly selected from the pool of remaining 214 candidates. All of the MHM searches were terminated after 300 DFT optimizations.

As summarized in Table 2, we found that none of the three major methods, i.e., AIRSS, USPEX, and MHM, can predict the

Table 2. Assessments of Polymer Structure Prediction Using Various Methods

search	optimizations	correct connectivity	correct structures
PSP	216	yes	yes
AIRSS	500	no	no
USPEX-1	1000	no	no
MHM-1	300	no	no
USPEX-2	180	yes	yes
MHM-2	3000	yes	yes

PE atomic connectivity if the initial structures/population is completely random. Only when proper constraints are imposed to enforce the atomic connectivity during the searches can the designated targets (the orthorhombic and monoclinic phases of PE) be predicted in USPEX-2 and MHM-2. Although USPEX-2 predicted the correct structures of PE after 9 generations, 29 generations were completed in order to fulfill the stopping criterion. In the case of MHM-2, two targets can only be recovered in two searches starting from the “right” initial structures, not in the other searches. Given that numerous parameters of an USPEX structure prediction could lead to some uncertainty in its efficiency, USPEX and the PSP can be considered to be efficiently comparable while MHM is less efficient. Unlike USPEX and MHM, a search using the PSP can be parallelized and scaled up very efficiently while requiring little human intervention to perform. Overall, the PSP is superior to AIRSS, USPEX, and MHM in predicting polymer crystal structures and thus is suitable for exploring the polymer space and generating polymer data in an autonomous and property-specific manner.

Nevertheless, it is important to note that the PSP has some known deficiencies. First, it depends heavily on the input polymer chain structure. For instance, because the chains of β -PVDF and δ -PVDF, shown in Figure 5, are so different; thus, the

PSP failed to predict one phase from the chain of the other. Second, for polymers whose atomic connectivity is the only known information, predicting the polymer chain conformation is highly nontrivial. One of the challenges is that the right chain conformation of a polymer, as shown in Figure 5 for two phases of PVDF, could be a bulk property; e.g., it is determined not only by the chain itself but also by the surrounding environment. Without this required input, the PSP cannot be used. Third, the energy difference between the polymer chain and crystal structures, as revealed by Figure 3, is typically very small ($\lesssim 1$ meV/atom), being comparable with the numerical errors of DFT calculations involving weak vdW interactions. Therefore, polymer structures may easily be misidentified if only energy is used for determining the structure stability. Finally, the PSP should be tested for complicated polymers with long and nonlinear side chain. These challenging problems are out of the scope of this work and could be the objective of a future studies.

In summary, we have developed an efficient method for computationally predicting polymer crystal structures from its chain model, which is a current great challenge in polymer informatics. Using this strategy, we have successfully recovered eight experimentally known structures of six linear polymers, including PE, PVDF, PVC, PPS, PAN, and ABPBO, and predicted new low-energy structures of β -PVDF, PAN, and PPS. Considering all of the distinct characteristics of polymers, the PSP is possibly the best approach to date for predicting polymer structures. The power of the PSP also lies in its simplicity, flexibility, and scalability; i.e., it can efficiently be used in modern high-performance and cloud computing facilities. Although determining the chain conformation remains a challenge, we believe that this method can be used for autonomously generating polymer data and exploring the polymer space in an efficient and comprehensive manner. Such a capability will be a critical component of polymer informatics ecosystems.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpclett.0c01553>.

CIF files (ZIP)

DFT optimized crystals structures recovered and discovered for PE, PVDF, PAN, ABPBO, PPS, and PVC and simulated X-ray diffraction patterns showing the difference between the DFT optimized recovered and discovered structures of β -PVDF, PAN, and PPS (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Tran Doan Huan — School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; orcid.org/0000-0002-8093-9426; Email: huan.tran@mse.gatech.edu

Rampi Ramprasad — School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States; orcid.org/0000-0003-4630-1565; Email: rampi.ramprasad@mse.gatech.edu

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpclett.0c01553>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is financially supported by Toyota Research Institute through the Accelerated Materials Design and Discovery program and Office of Naval Research through a Multi-University Research Initiative (MURI) grant (N00014-17-1-2656). The authors thank Vinit Sharma (ORNL and University of Tennessee) for useful discussions and XSEDE for computational support through Grant No. TG-DMR080058N.

■ REFERENCES

- (1) Ladd, M.; Palmer, R. *Structure Determination by X-ray Crystallography: Analysis by X-rays and Neutrons*; Springer US: Boston, MA, 2013; pp 111–159.
- (2) Oganov, A. R. Ed. *Modern Methods of Crystal Structure Prediction*; Wiley-VCH: Weinheim, Germany, 2011.
- (3) Mannodi-Kanakkithodi, A.; Treich, G.; Huan, T. D.; Ma, R.; Tefferi, M.; Cao, Y.; Sotzing, G.; Ramprasad, R. Rational co-design of polymer dielectrics for energy storage. *Adv. Mater.* **2016**, *28*, 6277–6291.
- (4) Zhu, Q.; Sharma, V.; Oganov, A. R.; Ramprasad, R. Predicting polymeric crystal structures by evolutionary algorithms. *J. Chem. Phys.* **2014**, *141*, 154102.
- (5) Sharma, V.; Wang, C. C.; Lorenzini, R. G.; Ma, R.; Zhu, Q.; Sinkovits, D. W.; Pilania, G.; Oganov, A. R.; Kumar, S.; Sotzing, G. A.; et al. Rational design of all organic polymer dielectrics. *Nat. Commun.* **2014**, *5*, 4845.
- (6) Baldwin, A. F.; Ma, R.; Mannodi-Kanakkithodi, A.; Huan, T. D.; Wang, C.; Marszalek, J. E.; Cakmak, M.; Cao, Y.; Ramprasad, R.; Sotzing, G. A. Poly(dimethyltin glutarate) as a prospective material for high dielectric applications. *Adv. Mater.* **2015**, *27*, 346.
- (7) Baldwin, A. F.; Huan, T. D.; Ma, R.; Mannodi-Kanakkithodi, A.; Tefferi, M.; Katz, N.; Cao, Y.; Ramprasad, R.; Sotzing, G. A. Rational design of organotin polyesters. *Macromolecules* **2015**, *48*, 2422.
- (8) Ma, R.; Sharma, V.; Baldwin, A. F.; Tefferi, M.; Offenbach, I.; Cakmak, M.; Weiss, R.; Cao, Y.; Ramprasad, R.; Sotzing, G. A. Rational design and synthesis of polythioureas as capacitor dielectrics. *J. Mater. Chem. A* **2015**, *3*, 14845–14852.
- (9) Huan, T. D.; Mannodi-Kanakkithodi, A.; Kim, C.; Sharma, V.; Pilania, G.; Ramprasad, R. A polymer dataset for accelerated property prediction and design. *Sci. Data* **2016**, *3*, 160012.
- (10) Glass, C. W.; Oganov, A. R.; Hansen, N. USPEX-Evolutionary crystal structure prediction. *Comput. Phys. Commun.* **2006**, *175*, 713–720.
- (11) Goedecker, S. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *J. Chem. Phys.* **2004**, *120*, 9911.
- (12) Amsler, M.; Goedecker, S. Crystal structure prediction using the minima hopping method. *J. Chem. Phys.* **2010**, *133*, 224104.
- (13) Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. CALYPSO: A method for crystal structure prediction. *Comput. Phys. Commun.* **2012**, *183*, 2063–2070.
- (14) Pickard, C. J.; Needs, R. J. Ab initio random structure searching. *J. Phys.: Condens. Matter* **2011**, *23*, 053201.
- (15) Oganov, A. R.; Pickard, C. J.; Zhu, Q.; Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **2019**, *4*, 331–348.
- (16) Ehrenstein, G. W. *Polymeric materials: structure, properties, applications*; Carl Hanser Verlag GmbH Co KG, 2012.
- (17) Huan, T. D.; Mannodi-Kanakkithodi, A.; Ramprasad, R. Accelerated materials property predictions and design using motif-based fingerprints. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2015**, *92*, 014106.
- (18) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Model.* **1988**, *28*, 31–36.
- (19) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864–B871.
- (20) Kohn, W.; Sham, L. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- (21) Peacock, A. *Handbook of Polyethylene: Structures: Properties, and Applications*, 1st ed.; CRC Press: New York, US, 2000.
- (22) Hasegawa, R.; Kobayashi, M.; Tadokoro, H. Molecular conformation and packing of poly(vinylidene fluoride). Stability of three crystalline forms and the effect of high pressure. *Polym. J.* **1972**, *3*, 591–599.
- (23) Lando, J. B.; Olf, H. G.; Peterlin, A. Nuclear magnetic resonance and x-ray determination of the structure of poly(vinylidene fluoride). *J. Polym. Sci., Part A-1: Polym. Chem.* **1966**, *4*, 941–951.
- (24) Wilkes, C. E.; Folt, V. L.; Krimm, S. Crystal Structure of Poly(vinyl chloride) Single Crystals. *Macromolecules* **1973**, *6*, 235–237.
- (25) Tabor, B. J.; Magré, E. P.; Boon, J. The crystal structure of poly-p-phenylene sulphide. *Eur. Polym. J.* **1971**, *7*, 1127–1133.
- (26) Hobson, R. J.; Windle, A. H. Crystalline structure of atactic polyacrylonitrile. *Macromolecules* **1993**, *26*, 6903–6907.
- (27) Fratini, A. V.; Cross, E. M.; O'Brien, J. F.; Adams, W. W. The structure of poly-2,5-benzoxazole (ABPBO) and poly-2,6-benzothiazole (ABPBT) fibers by X-ray diffraction. *J. Macromol. Sci., Part B: Phys.* **1985**, *24*, 159–179.
- (28) Kim, C.; Chandrasekaran, A.; Huan, T. D.; Das, D.; Ramprasad, R. Polymer Genome: A Data-Powered Polymer Informatics Platform for Property Predictions. *J. Phys. Chem. C* **2018**, *122*, 17575–17585.
- (29) Mannodi-Kanakkithodi, A.; Chandrasekaran, A.; Kim, C.; Huan, T. D.; Pilania, G.; Botu, V.; Ramprasad, R. Scoping the polymer genome: A roadmap for rational polymer dielectrics design and beyond. *Mater. Today* **2018**, *21*, 785–796.
- (30) Patra, A.; Batra, R.; Chandrasekaran, A.; Kim, C.; Huan, T. D.; Ramprasad, R. A multi-fidelity information-fusion approach to machine learn and predict polymer bandgap. *Comput. Mater. Sci.* **2020**, *172*, 109286.
- (31) Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *50*, 17953.
- (32) Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 558–561.
- (33) Kresse, G. Ab initio Molekular Dynamik für flüssige Metalle. *Ph.D. thesis*, Technische Universität Wien, 1993.
- (34) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.
- (35) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 11169–11186.
- (36) Liu, C.-S.; Pilania, G.; Wang, C.; Ramprasad, R. How Critical Are the van der Waals Interactions in Polymer Crystals? *J. Phys. Chem. A* **2012**, *116*, 9347–9352.
- (37) Lee, K.; Murray, E. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. Higher-accuracy van der Waals density functional. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010**, *82*, No. 081101(R).
- (38) Murray, E. D.; Lee, K.; Langreth, D. C. Investigation of exchange energy density functional accuracy for interacting molecules. *J. Chem. Theory Comput.* **2009**, *9*, 2754–2762.
- (39) Seto, T.; Hara, T.; Tanaka, K. Phase transformation and deformation processes in oriented polyethylene. *Jpn. J. Appl. Phys.* **1968**, *7*, 31.
- (40) Olsson, P. A. T.; Hyldgaard, P.; Schröder, E.; Jutemar, E. P.; Andreasson, E.; Kroon, M. Ab initio investigation of monoclinic phase stability and martensitic transformation in crystalline polyethylene. *Phys. Rev. Materials* **2018**, *2*, 075602.
- (41) Huan, T. D.; Boggs, S.; Teyssedre, G.; Laurent, C.; Cakmak, M.; Kumar, S.; Ramprasad, R. Advanced polymeric dielectrics for high energy density applications. *Prog. Mater. Sci.* **2016**, *83*, 236.
- (42) Chen, L.; Huan, T. D.; Ramprasad, R. Electronic Structure of Polyethylene: Role of Chemical, Morphological and Interfacial Complexity. *Sci. Rep.* **2017**, *7*, 6128.
- (43) Chen, L.; Tran, H. D.; Ramprasad, R. Atomistic mechanisms for chemical defects formation in polyethylene. *J. Chem. Phys.* **2018**, *149*, 234902.
- (44) Kumazoe, H.; Fukushima, S.; Tiwari, S.; Kim, C.; Huan, T. D.; Kalia, R. K.; Nakano, A.; Ramprasad, R.; Shimojo, F.; Vashishta, P. Hot-

Carrier Dynamics and Chemistry in Dielectric Polymers. *J. Phys. Chem. Lett.* **2019**, *10*, 3937–3943.

(45) Li, M.; Wondergem, H. J.; Spijkman, M.-J.; Asadi, K.; Katsouras, I.; Blom, P. W.; De Leeuw, D. M. Revisiting the δ -phase of poly(vinylidene fluoride) for solution-processed ferroelectric thin films. *Nat. Mater.* **2013**, *12*, 433–438.

(46) Huan, T. D.; Sharma, V.; Rossetti, G. A.; Ramprasad, R. Pathways towards ferroelectricity in hafnia. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *90*, 064111.

(47) Huan, T. D. Pressure-stabilized binary compounds of magnesium and silicon. *Phys. Rev. Materials* **2018**, *2*, 023803.