THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Crystal Prediction via Genetic Algorithms in a Model Chiral System

Published as part of The Journal of Physical Chemistry virtual special issue "Doros N. Theodorou Festschrift".

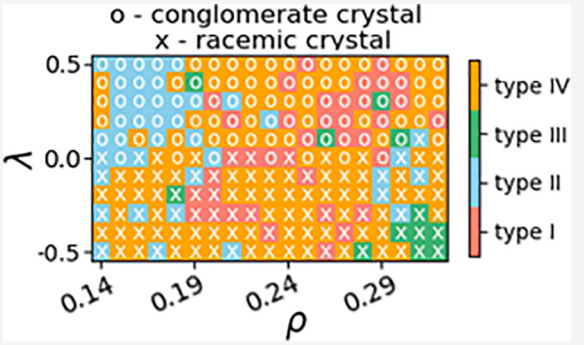
Nikolai D. Petsev,[#] Arash Nikoubashman,[#] Folarin Latinwo, Frank H. Stillinger, and Pablo G. Debenedetti*

Cite This: *J. Phys. Chem. B* 2022, 126, 7771−7780

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Chiral crystals and their constituent molecules play a prominent role in theories about the origin of biological homochirality and in drug discovery, design, and stability. Although the prediction and identification of stable chiral crystal structures is crucial for numerous technologies, including separation processes and polymorph selection and control, predictive ability is often complicated by a combination of many-body interactions and molecular complexity and handedness. In this work, we address these challenges by applying genetic algorithms to predict the ground-state crystal lattices formed by a chiral tetramer molecular model, which we have previously shown to exhibit complex fluid-phase behavior. Using this approach, we explore the relative stability and structures of the model's conglomerate and racemic crystals, and present a structural phase diagram for the stable Bravais crystal types in the zero-temperature limit.



I. INTRODUCTION

Crystals composed of chiral molecules (i.e., molecules that exist as isomers with nonsuperimposable mirror images) are important from both fundamental and technological perspectives.[1] Previous studies[2−4] indicate that the relative stability of conglomerate (chirally pure) and racemate crystals (composed of both enantiomers) provides a route toward the enrichment of one chiral entity relative to the other. From a fundamental standpoint, this offers a tantalizing proposal for the emergence of biological homochirality.[2−4] The chiral identity of biological molecules is essential to their function,[1,5,6] and hence, the appearance of biological homochirality is an outstanding question closely tied to the origin of terrestrial life. From a technological perspective, the ability to reliably predict and control chiral crystal structures can dramatically accelerate material discovery and design. This is particularly important for developing novel pharmaceuticals, since left-handed and right-handed enantiomers of the same drug molecule often exhibit drastically different physiological effects in the human body.[7−10] Therefore, effective separation of molecules based on chirality is often critical in industrial settings, and generating high-purity compounds by controlled crystallization is an auspicious alternative to the enantioselective asymmetric synthesis prevalent in the pharmaceutical industry.[11,12] Beyond new drug development and manufacture, crystal engineering (the selection of chemical substituents and processing conditions) as a route toward enantiopure chemical production is relevant to, for example, optoelectronic manufacturing,[13,14] catalysis,[15] and insecticides and agrochemicals.[16,17]

To address the need for tools that can offer molecular-level insight into processes such as chiral symmetry breaking and phase amplification, we previously developed a generic molecular model for studying chirality phenomena inspired by hydrogen peroxide,[18,19] the simplest chiral molecule; note, however, that this model system uses molecular parameters as tunable variables to explore the range of behaviors compatible with this type of chirality rather than as fitting parameters aimed at reproducing the properties of hydrogen peroxide or any other specific molecule. The chiral molecules consist of four sites, with chirality encoded in each molecule's dihedral angle; that is, bond constants are specified such that the tetrameric molecules can exist in either left-handed L- or right-handed D-configurations (see Figure 1). A critical component of this model is the coarse-grained parameter λ, a mean-field representation of short-range interactions that introduces an energetic preference for homochiral interactions when $\lambda > 0$ and heterochiral interactions when $\lambda < 0$. While simple in its formulation, the model's ability to tune the chiral bias has

Received: June 28, 2022
Revised: September 12, 2022
Published: September 26, 2022

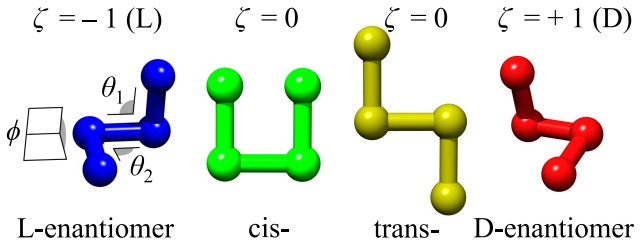


Figure 1. Visual representation of the chiral tetramer model. Molecules are able to adopt either left-handed L- (blue) or right-handed D- (red) enantiomer configurations. The chiral identity is measured by the order parameter ζ, whose value is shown in the figure above each corresponding configuration. For illustrative purposes, we include visualizations of the unstable cis- (green) and trans- (yellow) transition states.

enabled investigation into fluid-phase separation and chiral symmetry breaking in chiral systems through molecular dynamics (MD) simulations.[20−23] In particular, this model has been demonstrated to exhibit a wide range of complex chirality-induced phenomena in its various fluid phases, including phase amplification,[18] interconversion-controlled liquid−liquid phase separation,[20] and multiple fluid−fluid transitions.[23] The chiral tetramer model, in other words, has proven to be a remarkably versatile vehicle for exploring a wide range of complex chirality-induced fluid-phase behaviors, and in this work we focus on expanding our understanding of this model to the crystalline phase.

Despite the idealized nature of this model, however, it remains difficult to directly predict the range of stability of its conglomerate and racemate solid phases through application of plain MD calculations. This difficulty is not surprising, since direct crystallization of molecular systems from the liquid phase is well-known to be challenging in computer simulations due to the large number of potential crystal structures and molecular conformations.[24] The search for stable crystal structures is further complicated by the existence of polymorphic forms that differ only slightly in their free energy; for example, spherical particles on a face-centered cubic lattice have the exact same maximum packing fraction as in hexagonal close-packed structures, but exhibit a 10^{-3} k_BT larger free energy per sphere due to the different entropy of the void distribution.[25,26] Nevertheless, state-of-the-art advanced sampling techniques, such as the forward-flux sampling method,[27−29] umbrella sampling,[30,31] and crystal seeding methods,[32] as well as global optimization techniques such as basin-hopping,[33−35] evolutionary algorithms,[36−41] particle swarm optimization,[42,43] and simulated annealing,[44−47] have enabled a broad range of theoretical and computational investigations of the thermodynamics and kinetics of crystallization. Moreover, strides in crystal structure prediction leveraging highly refined empirical force fields and density functional theory (DFT) calculations have made possible the accurate prediction of optimal crystal structures formed by molecules with ever-increasing complexity, as illustrated by the series of Cambridge Crystallographic Data Center blind tests.[48] More recently, crystal structure prediction has shifted to include deep learning and data-driven approaches for structure identification, prediction, and design.[49−51] In this work, in contrast, we aim to isolate and highlight the effect of molecular chirality on the crystal phases of a model system that has already been shown to exhibit complex fluid-phase behavior.[18−20,23] Thus, we embark on the logical next step in the exploration of the rich behavior exhibited by our simple model system by investigating its possible crystal phases while minimizing computational expense and model degrees of freedom.

Given the success of evolutionary optimization techniques in crystal prediction,[38] in this work we utilize the framework of genetic algorithms (GAs) to systematically explore the possible solid crystalline phases of our tetramer model in the zero-temperature limit, and identify stable conglomerate and racemate crystals. GAs provide a biologically inspired stochastic route toward finding optimal parameters in high-dimensional space, which is a characteristic feature of many condensed matter physics systems.[38,52−54] This optimization strategy has been employed in the design of various soft materials, including granular and block copolymers,[55,56] DNA-grafted colloids,[57] multipolar colloids,[58,59] and nematic liquid crystals.[60] Other applications using GAs include the identification of stable crystals and polymorphs of small molecules,[61,62] structure prediction of interfaces in multicomponent systems,[63] and the docking of ligands in biomolecules.[64] Previous studies leveraging GAs for crystal prediction have primarily focused on ab initio DFT calculations for chemically specific systems,[65,66] with crystal lattice representation based on symmetry groups or standard crystal lattices.[67,68] Despite the intense research activity enabled by GAs, reports on the prediction of crystal structures comprised from chiral molecules are rather limited and focus on specific chemical substances (e.g., energetic materials[69] and plasmonic metasurfaces[70]). In contrast, this study instead seeks to utilize our simple, tunable model to extract chirality-specific phenomena that are general and transferrable to a broad spectrum systems.

The rest of this paper is organized as follows: In section IIA we briefly review the chiral tetramer model and its interaction potential, and in section IIB we describe the GA approach used in this work. In section III we present the results of the GA calculations, discuss the stable structures for conglomerates and racemates, and provide a structural phase diagram across a span of crystal densities. Finally, we offer concluding remarks and directions for future work in section IV.

II. MODEL AND METHODS

A. Chiral Tetramer Model. Our model system for generic chiral behavior consists of four-site molecules with two bond angles, θ_1 and θ_2, and a dihedral angle ϕ (see Figure 1).[18] Each site is equivalent with respect to intermolecular interactions. The relevant configurations a single molecule can adopt include the stable L- and D-enantiomer states, and unstable "trans"- and "cis"-transition states, also shown in Figure 1. The chiral order parameter ζ is a simple metric that quantifies the chiral state of an individual molecule, and can be defined as

$$\zeta(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = -\frac{\mathbf{r}_{12}\cdot(\mathbf{r}_{23}\times\mathbf{r}_{34})}{|\mathbf{r}_{12}||\mathbf{r}_{23}||\mathbf{r}_{34}|} \quad (1)$$

where the positions of the four monomers are given by the vectors $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, and $\mathbf{r}_4$, and relative vectors are defined as $\mathbf{r}_{12} \equiv \mathbf{r}_1 - \mathbf{r}_2$. In MD simulations of flexible and deformable molecules, the chiral order parameter varies continuously over the interval $\zeta \in [-1, 1]$,[18−20] but in the present work we treat the molecules as rigid bodies with fixed dihedral and bond angles in the two (L, D) mechanical equilibrium states with $\theta_1 = \theta_2 = \pi/2$, and $\phi = \pm\pi/2$. Thus, the chiral order parameter ζ^{α}

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

for molecule α can only adopt two possible values, that is, $\zeta^\alpha = -1$ if molecule α is in the left-handed L-configuration, and $\zeta^\alpha = +1$ if it is in the right-handed D-configuration (note that we use superscripts to index the tetramer molecules, while subscripts are used to index monomer sites on the molecules). The transition states ($\phi = 0$ and $\phi = \pi$) separating these two stable configurations lie on the hyperplane $\zeta^\alpha = 0$. The rigid body assumption is equivalently obtained by taking the positive-infinity limit of the intramolecular bond constants; future studies can relax this simplifaction to study the phonon spectra of the resulting crystal structures.

A key feature of our model is the ability to directly specify a chirality-dependent interaction potential between molecules (we ignore intramolecular interactions). The intermolecular potential energy function between monomer i from molecule α and monomer j from molecule γ is[18,19]

$$\Phi_{ij}^{\alpha\gamma}(\{\mathbf{r}_i^\alpha\}, \{\mathbf{r}_j^\gamma\}) = \varepsilon_0(1 + \lambda\zeta^\alpha\zeta^\gamma)\nu_{\mathrm{LJ}}\left(\frac{|\mathbf{r}_i^\alpha - \mathbf{r}_j^\gamma|}{\sigma_{tt}}\right) \tag{2}$$

Here, $\nu_{\mathrm{LJ}} = 4(\hat{r}^{-12} - \hat{r}^{-6})$ is the standard Lennard-Jones (LJ) potential with the energy prefactor factored out, $\hat{r} = |\mathbf{r}_j^\alpha - \mathbf{r}_k^\gamma|/\sigma_{tt}$ is the reduced monomer−monomer distance, ε_0 is the nominal interaction energy scale, and σ_{tt} is the monomer characteristic size. The "chiral renormalization parameter" λ in eq 2 allows us to tune the chiral bias of the system by giving energetic preference to locally enantiopure or racemic environments depending on its sign and magnitude. For example, if molecules α and γ have the same chiral identity, and hence, ζ^α and ζ^γ have the same sign, the molecular attraction becomes stronger when $\lambda > 0$, thus favoring homochiral interactions. Conversely, $\lambda < 0$ gives energetic preference for heterochiral interactions, or ones where tetramers α and γ have different chiral identities, that is, ζ^α and ζ^γ have opposite signs. Equation 2 reduces to the standard LJ potential for monomer pairs when there is no chiral bias, $\lambda = 0$. For our calculations, we use reduced LJ units such that ε_0 and the monomer diameter σ_{tt} are unity. The reduced bond length of the molecules is set to $b = 1.0583$, inspired by hydrogen peroxide (H_2O_2).[18,71−73] Note, however, that this model is not intended to accurately represent hydrogen peroxide, and this particular substance simply gives a basis for choosing qualitatively reasonable physical parameters for the tetramers.

The total potential energy $\Phi^{\alpha\gamma}$ for two interacting tetramers α and γ is given by summing all 16 pair interactions [eq 2] between monomers from molecule α, $i \in \alpha$, and monomers from molecule γ, $j \in \gamma$,

$$\Phi^{\alpha\gamma} = \sum_{i\in\alpha}\sum_{j\in\gamma,\gamma\neq\alpha}\Phi_{ij}^{\alpha\gamma} \tag{3}$$

Note that even though the potential energy in eq 3 is written as a pairwise sum over monomer sites, the energy for each interaction is rescaled by the chiral measures ζ^α and ζ^γ, which depend on the full configurations of both molecules α and γ. Hence, eq 2 represents an 8-body energy of interaction. For additional details on this interaction potential and its basic properties, we refer the reader to our previous work.[18−20]

In this study, we use lattice sums[74] to determine the potential energy per molecule, where our calculation is implemented similarly to the computation in ref 60. Using eq 3, we can write

$$\tilde{\phi}(X;\rho,\lambda) = \frac{1}{2}\sum_{\gamma\neq\alpha}\Phi^{\alpha\gamma} \tag{4}$$

where the sum goes over all molecules surrounding the molecule α at the origin that lie partially or fully within a cutoff radius $r_c = 4.0$. The precise number of molecules used for the lattice sum calculation depends on the choice of r_c and tetramer number density ρ, and we ensure that a sufficient number of periodic cells are generated in the x-, y-, and z-directions surrounding the molecule at the origin to give the required number of neighbors within r_c. We utilize a basis set composed of two molecules (here denoted by the "1" and "2" superscripts), which enables us to represent both conglomerate and racemic lattices. Here, $X = \{\mathbf{R}_i, \mathbf{b}_2, \mathbf{n}^1, \mathbf{n}^2, \zeta^1, \zeta^2\}$ denotes the Bravais lattice spanned by primitive vectors $\mathbf{R}_i \equiv \{\mathbf{a}_i\}$, with $i = 1, 2, 3$, $\mathbf{n}^1$ and $\mathbf{n}^2$ are orientation vectors for the molecules at the two basis points, respectively. The first basis vector is fixed to $\mathbf{b}_1 = (0,0,0)$, while the second basis vector $\mathbf{b}_2 = (\Delta x, \Delta y, \Delta z)$ is an optimizable parameter. Finally, ζ^1 and ζ^2 give the chiral identities of the molecules in the basis set and, hence, capture whether the crystal is conglomerate or racemic. To determine the ground state ($T = 0$) configurations of the crystal, we seek to minimize the potential energy $\tilde{\phi}$ [eq 4] at a particular chiral renormalization parameter λ and tetramer number density ρ with respect to the lattice parameters (see section IIB for additional discussion),

$$\phi(\rho,\lambda) = \min_{\{X\}}\tilde{\phi}(X;\rho,\lambda) \tag{5}$$

An exploration of crystal free energy minima at nonzero temperatures will be the subject of future investigation.

B. Brief Review of Genetic Algorithms for Crystal Systems. In this work, we leverage GAs, stochastic optimization algorithms inspired by evolutionary processes, to explore the high-dimensional energy landscape for crystals composed of rigid chiral molecules that interact through the intermolecular potential described in section IIA (eq 2). A single iteration of the GA search procedure consists of three parts: (*i*) encoding/decoding, (*ii*) crossover, and (*iii*) mutation. This section describes each of these steps in detail. For the (*i*) encoding/decoding step, each parameter required to specify a particular crystal lattice configuration is represented as a gene, enabling straightforward "genetic" transformations. The gene is composed of bits (a sequence of 0s and 1s) that represent the lattice parameters, and the number of bits used to encode these values specifies the numerical resolution for the optimization. In other words, using a high number of bits results in a fine-grained but often inefficient search process, whereas a small number of bits results in a coarse-grained representation of the energy landscape at the expense of numerical precision. The collection of genes that fully specifies a single crystal lattice is referred to as an individual I, and this "genotype" represents an instance of all parameters to be optimized. A group of individuals with the same genetic trajectory length is referred to as a generation. For a more comprehensive discussion of how the GA approach encodes information into bits, and performs transformations on the resulting genes, we refer the reader to ref 52.

The search procedure or genetic trajectory is induced in the (*ii*) crossover step, where the genes of "parent" individuals are randomly cut and recombined to create new genes that result

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

in a "child" individual. As a first step, the fitness of each individual I of the ith generation is calculated:

$$f_i(I) = \exp[-\phi(I)/\phi_0] \tag{6}$$

where $\phi(I)$ is the potential energy-per-molecule computed from the lattice sum eq 4, and ϕ_0 is an energy normalization factor that adjusts the scale of fitness values for numerical stability. The choice of ϕ_0 does not affect the final results, and in our work we use $\phi_0 = 10$, which narrows the range of possible values to $f(I) \sim [0, 100]$ for typical lattice energies. The probability that an individual I is selected for mating is then defined as (i.e., fitness proportionate selection or "roulette selection")

$$p(I) = \frac{f_i(I)}{\sum_I f_i(I)} \tag{7}$$

The subsequent $i + 1$ generation of individuals is then created by selecting parent individuals with a probability according to eq 7 and performing one-point crossover mating with the genetic material of the parents to generate a new genome for the child individual. The crossover point is determined by sampling from a uniform random distribution.

The final step in a single iteration of the GA search is (*iii*) mutation, where the genes of individuals are randomly altered to prevent the loss of genetic material. Hence, the mutation step introduces a stochastic element in the search that prevents individuals from being trapped in local minima. Steps (*ii*) and (*iii*) in the procedure outlined above are typically repeated for some fixed number of generations or iterated until convergence, based on an energy or fitness criterion. For a more detailed description of GAs, we refer the interested reader to previous reports on the subject.[39,52,60]

In the context of crystal systems, a careful selection of parameters is necessary to facilitate a rapid search for stable lattice configurations. Specifically, we assume that candidate crystal structures will be of the Bravais lattice type. With this assumption, a monatomic crystalline system in three dimensions can be defined by three primitive lattice vectors $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$, which equivalently can be parametrized using six parameters $\{a, x, y, \Theta, \Psi, \nu\}$. The lattice constant a is specified through the density of the crystal as $\rho^{-1} = |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|$ and therefore not included in the optimization procedure.[60] The parameters x and y are bound between $[0,1]$, and the lattice angles are defined on the ranges $\Theta \in [0, \pi/2]$, $\Psi \in [0, \pi]$, and $\nu \in [0, \pi/2]$. The transformation between lattice vectors and these parameters is given by

$$\begin{aligned}
\mathbf{a}_1 &= a(1, 0, 0),\\
\mathbf{a}_2 &= a(x\cos\nu, x\sin\nu, 0),\\
\mathbf{a}_3 &= a(xy\cos\Psi\cos\Theta, xy\sin\Psi\cos\Theta, xy\sin\Theta)
\end{aligned} \tag{8}$$

For monatomic crystals, an individual I is then defined as an instance of these parameters encoded into binary representation, $I_{\{a_i\}} = b_x b_y b_\Theta b_\Psi b_\nu$. The parametrization in eq 8 is not unique, and we include an additional constraint during optimization, requiring that the surface area of the unit cell is minimized.[60]

Compared to monatomic crystals, chiral systems may form enantiomerically pure (conglomerate) or racemate structures.[40] In addition, enantiomerically similar molecules may be distinct also on the basis of their orientation. Consequently, we implement the crystal search for our problem so that each lattice point represents the molecular center-of-mass, and we include one additional basis point to allow for exploring racemate structures. We found that including more basis points becomes prohibitively expensive from a computational standpoint. Each lattice point is also supplemented by two vectors $\mathbf{n}^1$ and $\mathbf{n}^2$, which describe the orientation of the two molecules. In our implementation, the chiral identity of the crystal is controlled by the measures ζ^1 and ζ^2 for the two basis points, and these identities are included as additional parameters (i.e., genes) to be optimized. In general, introducing more degrees of freedom in the lattice description appends additional bits to $I_{\{a_i\}}$, and hence, the final form of our genome is

$$I_{\{\mathbf{a}_i\}} = b_x b_y b_\Theta b_\Psi b_\nu b_{\Delta x} b_{\Delta y} b_{\Delta z} b_{n_x^1} b_{n_y^1} b_{n_z^1} b_{n_x^2} b_{n_y^2} b_{n_z^2} b_{\zeta^1} b_{\zeta^2} \tag{9}$$

Note that here we have included the displacement for the second basis point relative to the first, $\mathbf{b}_2 \equiv (\Delta x, \Delta y, \Delta z)$. Optimizing the set of genes in eq 9 is equivalent to the minimization in eq 5. Further details on the application of GAs to crystal systems and Bravais lattices, in particular, are provided in refs 52 and 60.

Lastly, to accelerate the convergence of the GA search, we implemented a variable resolution strategy, where the number of bits used to represent each individual is periodically changed every 10 generations by drawing a random number from a uniform distribution in the range $N_{\text{bit}} \in [7, 31]$ for all genes. We found that this "activated" search method significantly accelerates the optimization for our problem compared to traditional GAs with a fixed resolution (i.e., fixed number of bits). In between the periodic bit number changes, and also at the conclusion of the GA search, we perform a Hill-search optimization on the fittest individual I_i^* of the ith generation to remove errors due to the encoding resolution and ensure that parameters relax to the basin minimum. Finally, we apply "elitism" and transfer the most fit, Hill-search-optimized genotype I_i^* from generation i into the next generation $i + 1$, preserving the most fit individual's genetic information into the continuing search. Further improvements and a more quantitative exploration of the activated search GA strategy will be the subject of future work.

III. RESULTS AND DISCUSSION

We now apply the GA crystal search approach outlined in section IIB to optimize the fitness function in eq 6 and investigate the ground state crystal structures formed by the chiral tetramer model (section IIA).[75] In particular, we explore values for the coarse-grained chiral renormalization parameter λ in the range $\lambda \in [-0.5, 0.5]$ and molecular solid-phase densities $\rho \in [0.14, 0.32]$. The intermolecular potential in our calculations is given by eq 2, with values reported in reduced LJ units (see section IIA). In practice, a typical lattice in our calculations is composed of 128−432 molecules or 512−1728 monomers, depending on the choice of molecular number density. For a more comprehensive search of the high-dimensional parameter space, we initialize a collection of 57 independent GA walkers (i.e., independent initializations of the search algorithm that explore the energy landscape) for every (ρ, λ) point. The GA hyperparameters selected in the present study include a population composed of $N_I = 500$ individuals, optimization for a total of $N_G = 25000$ generations, and mutation probability $p_{\text{mutate}} = 0.05$ per bit.[60] After the GA search procedure completes, we select the lowest-energy structure from the resulting ensemble of GA trajectories.

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

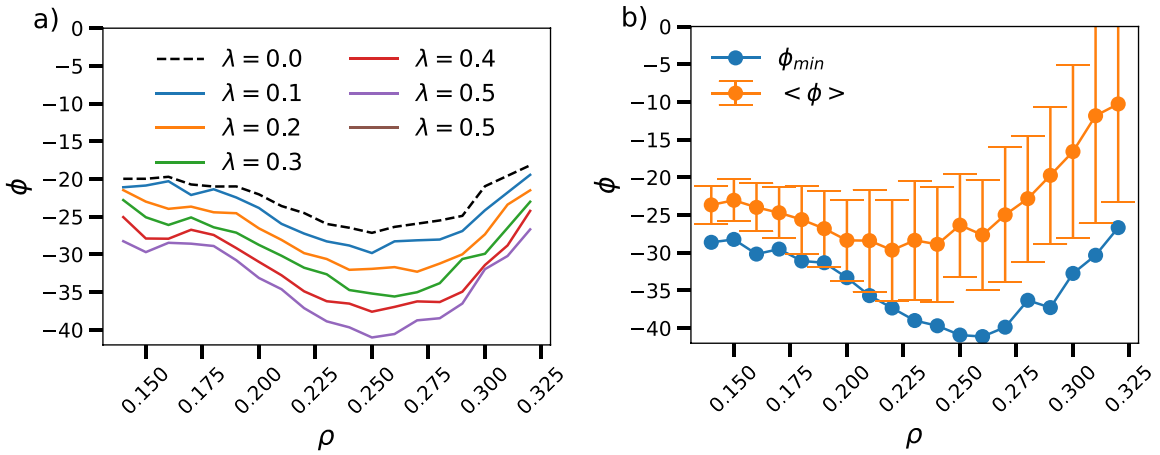


Figure 2. (a) Energy of most stable crystal configurations versus molecular number density for positive chiral bias $\lambda > 0$. Different colors denote different values of λ, as indicated. (b) The minimum energy ϕ_{min} for the $\lambda = 0.5$ case, compared to the average energy from the ensemble of 57 walkers, $\langle\phi\rangle$. Error bars indicate the standard deviation of energies in the ensemble.

First, we present results for the chiral renormalizations that favor enantiomerically pure crystal types, that is, $\lambda > 0$. Figure 2a shows the per-tetramer lattice energy for the most stable crystal structure as a function of molecular number density ρ and chiral renormalization in the interval $\lambda \in [0.0, 0.5]$. Different color curves correspond to different choices of λ, with $\lambda = 0$ (dashed black curve) giving the highest-energy structures for a given density overall, and $\lambda = 0.5$ (solid violet curve) giving the lowest energy ones. Note that the energy equation of state across different λ values possesses a generic feature with a minimum in the range of molecular densities $\rho \approx 0.24-0.26$, most frequently at $\rho = 0.25$, and exhibits increasing energy minimum depth for larger values of λ. Therefore, a larger chiral bias enables the crystal to adopt more energetically favorable configurations. Given that positive renormalization values favor locally homochiral interactions, we find that indeed most of the crystal structures obtained for $\lambda > 0$ are enantiopure (conglomerate crystals).

Interestingly, we identified one exception at $\lambda = 0.1$ and $\rho = 0.31$, where a racemate simple orthorhombic crystal with energy $\phi = -21.91$ was observed, despite having a positive chiral renormalization parameter. After revisiting this particular state point and performing another ensemble of GA optimization at identical conditions but different seeds for the random number generator, we obtained a conglomerate structure with slightly higher energy, $\phi = -20.30$. This structure is also a simple orthorhombic lattice, and differs only in chiral character. We also observed such competing crystal structures with comparable energies but different chiral identities in some our exploratory test runs at higher densities (i.e., $\rho > 0.30$) and near vanishing renormalizations $\lambda \to 0^+$ (results not shown), though these structures were ultimately replaced by more stable conglomerate ones when collecting our final set of data. Finally, we note that the probability of observing this metastability diminishes drastically with increasing λ, and we observed no racemate crystals when $\lambda > 0.1$ in our final calculations.

To illustrate the variance in energy across the ensemble of GA optimized structures, we focus on the $\lambda = 0.5$ case in Figure 2b and show the minimum energy curve $\phi(\rho, \lambda = 0.5)$ compared against the average energy of the ensemble for every point $\langle\phi\rangle$ $(\rho, \lambda = 0.5)$. Error bars for $\langle\phi\rangle$ represent the standard deviation in energies across the 57 independent trajectories that comprise the ensemble. The error bars for curves corresponding to the $\lambda \neq 0.5$ cases look similar to the representative case in Figure 2b and are, therefore, omitted in Figure 2a for clarity. Note that the standard deviation of the ensemble energy becomes larger with increasing ρ, illustrating that it is more challenging for the GA algorithm to converge at higher densities and highlighting the need for the ensemble approach utilized in this paper.

The most stable conglomerate crystal found by the GA was at the point $\lambda = 0.5$ and $\rho = 0.25$, with a per-tetramer lattice energy of $\phi = -40.97$. This is the lowest energy crystal lattice obtained from our entire study. A visualization of this particular structure is presented in Figure 3, where a single lattice cell comprised of two tetramer molecules is highlighted. In order to determine the identity of the Bravais lattice obtained from this search, we applied the NIST*LATTICE classification strategy[76,77] to the GA optimized primitive vectors, combined with manual categorization for important state points. This algorithm is based on cell type matching using a reduced form of the six scalar products between the

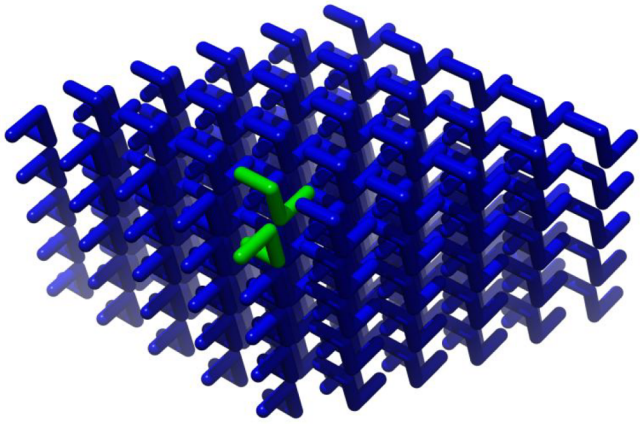


Figure 3. Visualization for the most stable conglomerate crystal structure for $\lambda > 0$, which occurred at the point $(\rho, \lambda) = (0.25, 0.5)$. Here, all molecules are in the L-configuration. Two molecules from a single unit cell are colored in green to aid the eye. The two green molecules represent a single cell of the lattice that is repeated to generate the full crystal structure.

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

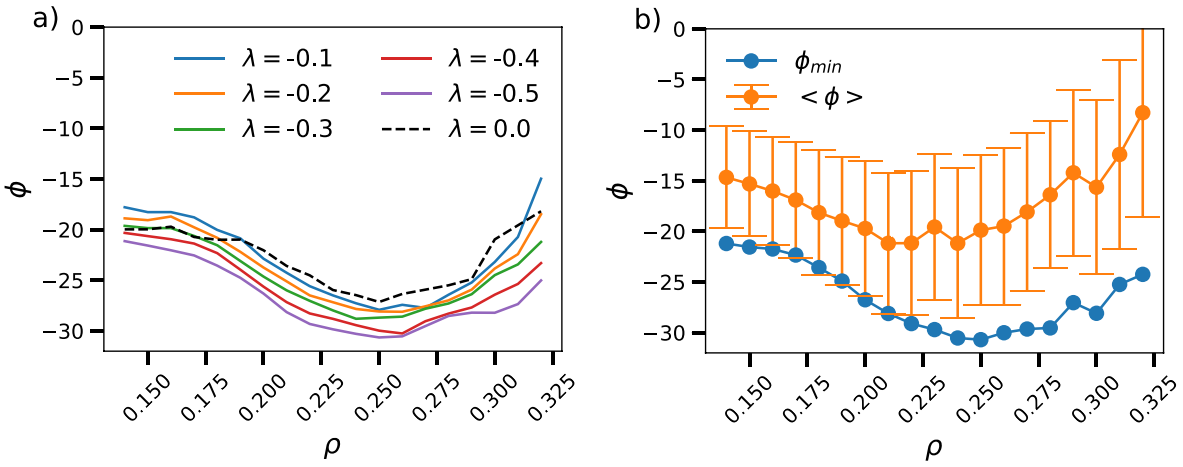


Figure 4. (a) Energy vs molecular number density ρ of the most stable crystal configurations for negative chiral bias, $\lambda < 0$. Different colors denote different choices for the chiral renormalization parameter λ, as indicated. (b) Energies of the most stable racemic structures obtained from our calculations, ϕ_{min}, compared to the ensemble average $\langle\phi\rangle$.

three primitive lattice vectors, $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$. Due to the numerical precision of the GA, the optimized lattice vectors will inevitably deviate somewhat from the ones of perfect Bravais crystals, thus necessitating some tolerance in the identification procedure; we chose a relative tolerance of 10^{-4} and absolute tolerance of 0.15 of the magnitude of the largest primitive vector. We note that this approach does not unambiguously identify lattices, and the specific choice of absolute tolerance impacts the classification. In particular, the distinction between type II and type IV lattices (described below) is affected by the choice of absolute tolerance, whereas other lattice types are less sensitive. Nevertheless, the present choice of these parameters does identify clear patterns in the data. Our classification algorithm identified a total of 13 distinct lattices, which can be broadly grouped into four general lattice types based on the primitive vector lengths $a_1 = |\mathbf{a}_1|$, $a_2 = |\mathbf{a}_2|$, and $a_3 = |\mathbf{a}_3|$: (*i*) type I lattices, where the three primitive vectors obey $a_1 = a_2 = a_3$, (*ii*) type II lattices, in which case $a_1 = a_2$, (*iii*) type III lattices, where $a_2 = a_3$, and (*iv*) type IV lattices where $a_1 \leq a_2 \leq a_3$, which captures lattices that did not belong to types I−III.

From this analysis, we determined that the global energy minimum at $(\rho, \lambda) = (0.25, 0.5)$ corresponds to a conglomerate rhombohedral lattice, which has a significantly lower energy than our previously reported conglomerate crystal lattice [$\phi = -29.4$ for $(\rho, \lambda) = (0.25, 0.5)$];[18] the structure presented in our previous work was obtained via a combination of inherent structure calculations for a pair of tetramer molecules, geometrical considerations, and low finite temperature isobaric molecular dynamics calculations. At this state point where the most stable structure obtained, the GA search identified a collection of eight unique structures among the 57 independent runs, where a triclinic lattice was the most dominant Bravais lattice observed (44% of structures), followed by oblique rhombic prism (21%) and rhombohedral (12%) lattices. The remaining observed structures included C-centered monoclinic, body-centered monoclinic, face-centered orthorhombic, and body-centered tetragonal lattices.

Next, we turn our attention to renormalizations that favor racemic crystal types, $\lambda < 0$. Figure 4a shows the lattice energy per tetramer for the most stable crystal structures as a function of molecular number density ρ over a range of negative renormalization parameter values. We also provide in Figure 4b the ensemble average of the energy and corresponding standard deviation for the most stable crystal structure (in this case, $\lambda = -0.5$). Note that for all of the $\lambda < 0$ genetic trajectories, only racemate structures were obtained, and hence the affinity for local heterochirality imposed by negative λ values appears to always produce racemates. This behavior is in contrast to positive renormalizations, where our calculations still produced a heterochiral structure at $\lambda = 0.1$ and $\rho = 0.31$. Similar to the conglomerate case, the lattice energy per tetramer for racemates exhibits a general form with a minimum in the range of $\rho \approx 0.24-0.26$. Moreover, there is again a linear relationship between the minimum lattice energy per tetramer ϕ_{min} and the chiral bias λ (see discussion below). The most stable racemate found from our application of the GA is at $(\rho, \lambda) = 0.25, -0.5)$ with a lattice energy of $\phi_{min} = -30.39$. The crystal structure for the most stable racemate crystal is a primitive oblique rhombic prism. A total of eight unique structures were identified across the full ensemble at this particular state point, with the most common structure being a triclinic lattice (37% of structures), followed by oblique rhombic prism (21% of structures), and rhombohedral hR (18% of structures) lattices. Other lattice types observed in the ensemble included C-centered monoclinic, body-centered monoclinic, rhombohedral, and body-centered tetragonal. A snapshot of the lattice corresponding to the most stable racemic crystal is provided in Figure 5. The full energy surface determined from our calculations is summarized in Figure 6, and the crystal structure, density, and chiral identity for the most stable lattice at every λ value are summarized in Table 1.

We now focus on the differences between the conglomerate and racemate crystal predictions. Figure 7 shows the minimum lattice energy per tetramer ϕ_{min} across all investigated ρ values as a function of renormalization λ. Both conglomerates and racemates exhibit a linear relationship between ϕ_{min} and λ, but the dependence is stronger for conglomerates than racemates, as evidenced by the difference in slopes shown in Figure 7. Given the form of the potential energy function of our tetramer model, this result can be explained by drawing a rough analogy to the Ising model with up to second nearest neighbor interactions.[19,78,79] Of course, there are subtle differences between the two models such as the presence of more than second nearest neighbor interactions and molecularity in our chiral tetramer model. Nevertheless, the conglomerate crystals

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

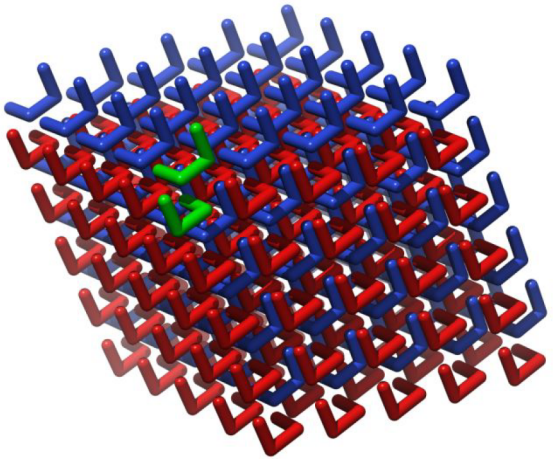


Figure 5. Visualization of the most stable racemic crystal lattice, which occurs at the point $(\rho, \lambda) = (0.25, -0.5)$. Here, red denotes molecules in the D-configuration, and blue denotes molecules in the L-configuration. We have highlighted in green two molecules from a single lattice cell to aid visualization of what a single unit cell looks like.

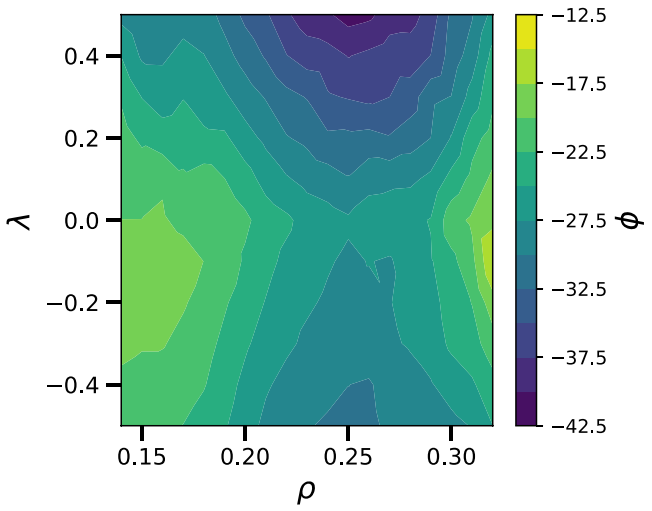


Figure 6. Most stable crystal energy as a function of both crystal density ρ and chiral renormalization λ.

Table 1. Summary of Most Stable Crystal Structure for Every Chiral Renormalization Considered in Our Study[a]

λ	ρ	chirality	structure
−0.5	0.25	racemate	oblique rhombic prism (primitive)
−0.4	0.26	racemate	monoclinic/triclinic (primitive)
−0.3	0.24	racemate	monoclinic (primitive)
−0.2	0.26	racemate	oblique rhombic prism (primitive)
−0.1	0.25	racemate	tetratgonal (body-centered)
0.0	0.25	conglomerate	triclinic (primitive)
0.1	0.25	conglomerate	triclinic (primitive)
0.2	0.27	conglomerate	tetragonal (body-centered)
0.3	0.26	conglomerate	tetragonal (body-centered)
0.4	0.25	conglomerate	triclinic (primitive)
0.5	0.25	conglomerate	rhombohedral

[a]The table summarizes the density at which the most stable crystal was determined, as well as the crystal chirality and its structure.

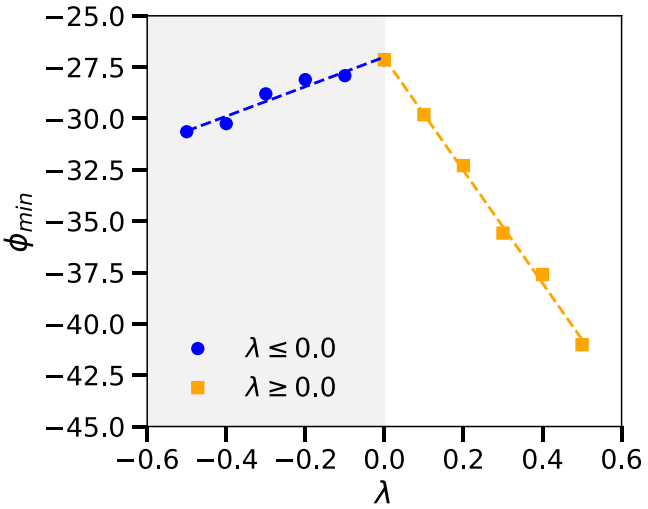


Figure 7. Energy of the most stable crystal, ϕ_{min}, across all densities as a function of chiral renormalization (markers), with linear fits to the data (dashed line).

have similar features to energetically favored ferromagnetic parallel spins, while racemates are analogous to antiferromagnetic systems that favor antiparallel spins. In this context, increasing the magnitude of λ corresponds to increasing the coupling constant in the Hamiltonian of the Ising model. Note that unfavorable second nearest neighbor chiral interactions, which can lead to cancellation effects, are present in racemates but are absent in conglomerates. Specifically, the second nearest neighbor interactions will diminish the effect of the coupling constant λ for racemates with $\lambda < 0$, but will increase for conglomerates with $\lambda > 0$. Therefore, an increase in the magnitude of λ will produce a less dramatic effect for favorable racemates when enantiomers (antiparallel spins) are nearest neighbors, than for conglomerates when chirally equivalent molecules (parallel spins) are nearest neighbors. At low densities, where a small number of molecules contribute to the total energy, the presence of energetically unfavorable second nearest neighbor interactions can potentially give rise to unexpected results; for example, the bias-free case gives lower-energy structures compared to $\lambda < 0$ at low densities, $\rho < 0.19$ (see Figure 4a).

We conclude our discussion of the GA crystal predictions for the chiral tetramer model by computing a structure phase diagram (Figure 8). While there is some statistical noise in the results, this diagram illustrates a clear preference for type II lattices at lower densities $\rho \leq 0.18$ and $\lambda > 0$, type III lattices at high densities $\rho > 0.29$ and $\lambda < 0$, and a mix of types I and IV lattices at intermediate densities ($0.19 \leq \rho \leq 0.28$; Figure 8). As described earlier, structures are predominantly conglomerate when $\lambda > 0$, and racemic for $\lambda < 0$. Along $\lambda = 0$, there is no preference for conglomerate or racemic structures, and hence both types are observed. There is no obvious strong link between crystal chiral identity and lattice structure, though at lower densities, type II structures were more frequently observed when $\lambda > 0$, and at high densities, type III structures predominantly appear when $\lambda < 0$. The most common lattice type observed across the full span of λ and ρ values is primitive triclinic, which comprised 33% of all the observed lattices. Oblique rhombic prism structures accounted for 32%, C-centered monoclinic 11%, and rhombohedral 10%. The remaining observed lattice structures included body-centered

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

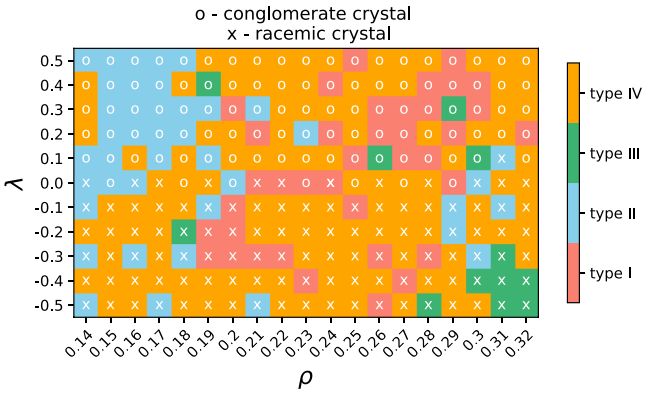


Figure 8. Summary of most stable crystal type across all molecular number densities ρ and chiral bias λ considered in this study. For a clearer presentation, we classify crystals into the four lattice types, as denoted by the four main colors (orange, green, blue, red). The "o" marker denotes that the crystal at that point is conglomerate, and the "x" marker denotes racemic crystals.

monoclinic, C-centered, face-centered, and body-centered orthorhombic, primitive and body-centered tetragonal, and body-centered cubic. Each of these lattice types accounted for <10% of all structures.

IV. CONCLUSION

Genetic algorithms represent a powerful class of tools for predicting and analyzing ground state crystal structures, and we apply this framework to chiral systems and predict conglomerate (chirally pure) and racemate (equal part enantiomers) crystals over a broad range of conditions. After identifying stable crystals with respect to molecular number density and chiral renormalization parameter, we applied the NIST*LATTICE approach to classify the optimized Bravais lattices computed by the genetic algorithm. This approach enabled us to develop a phase diagram for the chiral tetramer model, classifying stable structures based on both lattice type and chirality. Moreover, our findings revealed the presence of racemic crystals for small positive chiral renormalization values, suggesting the possibility of metastable states that differ only in chiral identity. This result has tantalizing prospects for better understanding the kinetics of crystallization and the mechanistic pathways when a metastable phase has a different chiral identity from its thermodynamically stable phase. In addition, the chirality-induced metastability present in our model may play a role in low temperature amorphous states, and may pose new questions relevant to glass physics or the recently reported critical point confluence phenomenon.[80]

Several aspects of this present work remain open for further exploration. First, from an algorithmic standpoint, the efficiency of the activated and expanded-ensemble genetic algorithm search presented here may be improved through further study, and by leveraging more sophisticated heuristics. In addition, advanced sampling techniques for rate calculations and the identification of mechanistic pathways of the crystallization of the tetramer model may elucidate the effects of the solid-phase chiral metastability. The chiral tetramer model has already been implemented into the LAMMPS simulation package,[19,81,82] and further investigation of the chiral tetramer model solid-phase behavior can be explored using molecular dynamics simulations using the crystal structures obtained in this work. Such simulations would explore the phonon spectra of these lattices, and can also open a pathway for studying the effect of strong compression and high crystal densities, where pressure-induced chirality inversion may occur.[18] Finally, we anticipate that the investigation of the additives and interactions that stabilize conglomerates and racemates will be of great interest in crystal design and engineering.

■ ASSOCIATED CONTENT

Data Availability Statement

The source code for the genetic algorithm optimization of crystal structures formed by the chiral tetramer molecular model described in this study is openly available for download in the Princeton DataSpace repository at 10.34770/CHAG-5F03.[75]

■ AUTHOR INFORMATION

Corresponding Author

Pablo G. Debenedetti − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0003-1881-1728; Email: pdebene@princeton.edu

Authors

Nikolai D. Petsev − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0002-8647-720X

Arash Nikoubashman − *Institute of Physics, Johannes Gutenberg University Mainz, 55128 Mainz, Germany;* orcid.org/0000-0003-0563-825X

Folarin Latinwo − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States; Synopsys Inc., Austin, Texas 78746, United States;* orcid.org/0000-0002-2331-8607

Frank H. Stillinger − *Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0002-1225-8186

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.2c04501

Author Contributions

#N.D.P. and A.N. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

A.N. acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Project Nos. 274340645 and 470113688. P.G.D. acknowledges the support of the National Science Foundation (Award CHE-1856704). The simulations presented in this article were performed on computational resources managed and supported by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of Information Technology's High Performance Computing Center and Visualization Laboratory at Princeton University. Molecular visualization was performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS No. P41-GM103311).

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

REFERENCES

(1) Yus, M.; Guijarro, A. *The Origin of Chirality in the Molecules of Life: A Revision from Awareness to the Current Theories and Perspectives of This Unsolved Problem*, 1st ed.; Royal Society of Chemistry: Cambridge, 2009.

(2) Blackmond, D. G. The Origin of Biological Homochirality. *Cold Spring Harb. Perspect. Biol.* **2019**, *11* (3), a032540.

(3) Viedma, C. Enantiomeric Crystallization from DL-Aspartic and DL-Glutamic Acids: Implications for Biomolecular Chirality in the Origin of Life. *Orig. Life Evol. Biosph.* **2001**, *31* (6), 501−509.

(4) Lombardo, T. G.; Stillinger, F. H.; Debenedetti, P. G. Thermodynamic Mechanism for Solution Phase Chiral Amplification via a Lattice Model. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (36), 15131−15135.

(5) Lough, W. J.; Wainer, I. W. *Chirality in Natural and Applied Science*; Wiley-Blackwell: Oxford, 2006.

(6) Hoehlig, K.; Bethge, L.; Klussmann, S. Stereospecificity of Oligonucleotide Interactions Revisited: No Evidence for Heterochiral Hybridization and Ribozyme/DNAzyme Activity. *PLoS One* **2015**, *10* (2), No. e0115328.

(7) Teo, S. K.; Colburn, W. A.; Tracewell, W. G.; Kook, K. A.; Stirling, D. I.; Jaworsky, M. S.; Scheffler, M. A.; Thomas, S. D.; Laskin, O. L. Clinical Pharmacokinetics of Thalidomide. *Clin. Pharmacokinet.* **2004**, *43* (5), 311−327.

(8) Kean, W. F.; Howard-Lock, H. E.; Lock, C. J. L. Chirality in Antirheumatic Drugs. *Lancet* **1991**, *338* (8782), 1565−1568.

(9) Chhabra, N.; Aseri, M. L.; Padmanabhan, D. A Review of Drug Isomerism and Its Significance. *Int. J. Appl. Basic Med. Res.* **2013**, *3* (1), 16−18.

(10) *Chiral Drugs: Chemistry and Biological Action*, 1st ed.; Lin, G.-Q., You, Q.-D., Cheng, J.-F., Eds.; Wiley: Hoboken, N.J., 2011.

(11) Lorenz, H.; Seidel-Morgenstern, A. Processes To Separate Enantiomers. *Angew. Chem., Int. Ed.* **2014**, *53* (5), 1218−1250.

(12) Cao, M.; Monson, P. A. A Study of the Phase Behavior of a Simple Model of Chiral Molecules and Enantiomeric Mixtures. *J. Chem. Phys.* **2005**, *122* (5), 054505.

(13) Nugent, W. A.; RajanBabu, T. V.; Burk, M. J. Beyond Nature's Chiral Pool: Enantioselective Catalysis in Industry. *Science* **1993**, *259* (5094), 479−483.

(14) Govan, J.; Gun'ko, Y. K. Recent Progress in Chiral Inorganic Nanostructures. *Nanoscience* **2016**, *3*, 1−30.

(15) Mallat, T.; Orglmeister, E.; Baiker, A. Asymmetric Catalysis at Chiral Metal Surfaces. *Chem. Rev.* **2007**, *107* (11), 4863−4890.

(16) Liu, W.; Gan, J.; Schlenk, D.; Jury, W. A. Enantioselectivity in Environmental Safety of Current Chiral Insecticides. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (3), 701−706.

(17) Aguiar, A. R.; Alvarenga, E. S.; Silva, E. M.; Farias, E. S.; Picanço, M. C. Synthesis, Insecticidal Activity, and Phytotoxicity of Novel Chiral Amides. *Pest Manag. Sci.* **2019**, *75* (6), 1689−1696.

(18) Latinwo, F.; Stillinger, F. H.; Debenedetti, P. G. Molecular Model for Chirality Phenomena. *J. Chem. Phys.* **2016**, *145* (15), 154503.

(19) Petsev, N. D.; Stillinger, F. H.; Debenedetti, P. G. Effect of Configuration-Dependent Multi-Body Forces on Interconversion Kinetics of a Chiral Tetramer Model. *J. Chem. Phys.* **2021**, *155* (8), 084105.

(20) Uralcan, B.; Longo, T. J.; Anisimov, M. A.; Stillinger, F. H.; Debenedetti, P. G. Interconversion-Controlled Liquid-Liquid Phase Separation in a Molecular Chiral Model. *J. Chem. Phys.* **2021**, *155* (20), 204502.

(21) Longo, T. J.; Anisimov, M. A. Phase Transitions Affected by Natural and Forceful Molecular Interconversion. *J. Chem. Phys.* **2022**, *156* (8), 084502.

(22) Longo, T. J.; Shumovskyi, N. A.; Asadov, S. M.; Buldryev, S. V.; Anisimov, M. A. Structure Factor of a Phase Separating Binary Mixture with Natural and Forceful Interconversion of Species. *J. Non-Cryst. Solids X* **2022**, *13*, 100082.

(23) Wang, Y.; Stillinger, F. H.; Debenedetti, P. G. Fluid-Fluid Phase Transitions in a Chiral Molecular Model. *J. Chem. Phys.* **2022**, *157* (8), 084501.

(24) Martoňák, R.; Laio, A.; Parrinello, M. Predicting Crystal Structures: The Parrinello-Rahman Method Revisited. *Phys. Rev. Lett.* **2003**, *90* (7), 075503.

(25) Bolhuis, P. G.; Frenkel, D.; Mau, S.-C.; Huse, D. A. Entropy Difference between Crystal Phases. *Nature* **1997**, *388* (6639), 235−236.

(26) Mahynski, N. A.; Panagiotopoulos, A. Z.; Meng, D.; Kumar, S. K. Stabilizing Colloidal Crystals by Leveraging Void Distributions. *Nat. Commun.* **2014**, *5* (1), 4472.

(27) Allen, R. J.; Warren, P. B.; ten Wolde, P. R. Sampling Rare Switching Events in Biochemical Networks. *Phys. Rev. Lett.* **2005**, *94* (1), 018104.

(28) Allen, R. J.; Frenkel, D.; ten Wolde, P. R. Simulating Rare Events in Equilibrium or Nonequilibrium Stochastic Systems. *J. Chem. Phys.* **2006**, *124* (2), 024102.

(29) Haji-Akbari, A.; Debenedetti, P. G. Direct Calculation of Ice Homogeneous Nucleation Rate for a Molecular Model of Water. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (34), 10582−10588.

(30) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187−199.

(31) Reinhardt, A.; Doye, J. P. K.; Noya, E. G.; Vega, C. Local Order Parameters for Use in Driving Homogeneous Ice Nucleation with All-Atom Models of Water. *J. Chem. Phys.* **2012**, *137* (19), 194504.

(32) Espinosa, J. R.; Vega, C.; Valeriani, C.; Sanz, E. Seeding Approach to Crystal Nucleation. *J. Chem. Phys.* **2016**, *144* (3), 034501.

(33) Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111−5116.

(34) Wales, D. J.; Scheraga, H. A. Global Optimization of Clusters, Crystals, and Biomolecules. *Science* **1999**, *285* (5432), 1368−1372.

(35) Amsler, M.; Goedecker, S. Crystal Structure Prediction Using the Minima Hopping Method. *J. Chem. Phys.* **2010**, *133* (22), 224104.

(36) Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*, 13th ed.; Addison-Wesley Professional, 1989.

(37) Ashlock, D. *Evolutionary Computation for Modeling and Optimization*, 2006th ed.; Springer, 2006.

(38) Oganov, A. R.; Glass, C. W. Crystal Structure Prediction Using Ab Initio Evolutionary Techniques: Principles and Applications. *J. Chem. Phys.* **2006**, *124* (24), 244704.

(39) Oganov, A. R.; Glass, C. W. Evolutionary Crystal Structure Prediction as a Tool in Materials Design. *J. Phys.: Condens. Matter* **2008**, *20* (6), 064210.

(40) Oganov, A. R.; Ma, Y.; Lyakhov, A. O.; Valle, M.; Gatti, C. Evolutionary Crystal Structure Prediction as a Method for the Discovery of Minerals and Materials. *Rev. Mineral. Geochem.* **2010**, *71* (1), 271−298.

(41) Oganov, A. R.; Lyakhov, A. O.; Valle, M. How Evolutionary Crystal Structure Prediction Works−and Why. *Acc. Chem. Res.* **2011**, *44*, 227−237.

(42) Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. Crystal Structure Prediction via Particle-Swarm Optimization. *Phys. Rev. B* **2010**, *82* (9), 094116.

(43) Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. CALYPSO: A Method for Crystal Structure Prediction. *Comput. Phys. Commun.* **2012**, *183* (10), 2063−2070.

(44) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. Optimization by Simulated Annealing. *Science* **1983**, *220* (4598), 671−680.

(45) Kirkpatrick, S. Optimization by Simulated Annealing: Quantitative Studies. *J. Stat. Phys.* **1984**, *34* (5), 975−986.

(46) Pannetier, J.; Bassas-Alsina, J.; Rodriguez-Carvajal, J.; Caignaert, V. Prediction of Crystal Structures from Crystal Chemistry Rules by Simulated Annealing. *Nature* **1990**, *346* (6282), 343−345.

https://doi.org/10.1021/acs.jpcb.2c04501
J. Phys. Chem. B 2022, 126, 7771−7780

(47) Earl, D. J.; Deem, M. W. Parallel Tempering: Theory, Applications, and New Perspectives. *Phys. Chem. Chem. Phys.* **2005**, *7* (23), 3910−3916.
(48) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; Case, D. H.; Chadha, R.; Cole, J. C.; Cosburn, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; DiStasio, R. A., Jr; Dzyabchenko, A.; van Eijck, B. P.; Elking, D. M.; van den Ende, J. A.; Facelli, J. C.; Ferraro, M. B.; Fusti-Molnar, L.; Gatsiou, C.-A.; Gee, T. S.; de Gelder, R.; Ghiringhelli, L. M.; Goto, H.; Grimme, S.; Guo, R.; Hofmann, D. W. M.; Hoja, J.; Hylton, R. K.; Iuzzolino, L.; Jankiewicz, W.; de Jong, D. T.; Kendrick, J.; de Klerk, N. J. J.; Ko, H.-Y.; Kuleshova, L. N.; Li, X.; Lohani, S.; Leusen, F. J. J.; Lund, A. M.; Lv, J.; Ma, Y.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Needs, R. J.; Neumann, M. A.; Nyman, J.; Obata, S.; Oberhofer, H.; Oganov, A. R.; Orendt, A. M.; Pagola, G. I.; Pantelides, C. C.; Pickard, C. J.; Podeszwa, R.; Price, L. S.; Price, S. L.; Pulido, A.; Read, M. G.; Reuter, K.; Schneider, E.; Schober, C.; Shields, G. P.; Singh, P.; Sugden, I. J.; Szalewicz, K.; Taylor, C. R.; Tkatchenko, A.; Tuckerman, M. E.; Vacarro, F.; Vasileiadis, M.; Vazquez-Mayagoitia, A.; Vogt, L.; Wang, Y.; Watson, R. E.; de Wijs, G. A.; Yang, J.; Zhu, Q.; Groom, C. R. Report on the Sixth Blind Test of Organic Crystal Structure Prediction Methods. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72* (4), 439−459.
(49) Ryan, K.; Lengyel, J.; Shatruk, M. Crystal Structure Prediction via Deep Learning. *J. Am. Chem. Soc.* **2018**, *140* (32), 10158−10168.
(50) Kim, S.; Noh, J.; Gu, G. H.; Aspuru-Guzik, A.; Jung, Y. Generative Adversarial Networks for Crystal Structure Prediction. *ACS Central Science* **2020**, *6*, 1412−1420.
(51) Wengert, S.; Csányi, G.; Reuter, K. T.; Margraf, J. Data-Efficient Machine Learning for Molecular Crystal Structure Prediction. *Chem. Sci.* **2021**, *12* (12), 4536−4546.
(52) Gottwald, D.; Kahl, G.; Likos, C. N. Predicting Equilibrium Structures in Freezing Processes. *J. Chem. Phys.* **2005**, *122* (20), 204503.
(53) van de Walle, A. Genesis of Crystal Structures. *Nat. Mater.* **2005**, *4* (5), 362−363.
(54) Patra, T. K.; Meenakshisundaram, V.; Hung, J.-H.; Simmons, D. S. Neural-Network-Biased Genetic Algorithms for Materials Design: Evolutionary Algorithms That Learn. *ACS Comb. Sci.* **2017**, *19* (2), 96−107.
(55) Jaeger, H. M.; de Pablo, J. J. Perspective: Evolutionary Design of Granular Media and Block Copolymer Patterns. *APL Mater.* **2016**, *4* (5), 053209.
(56) Khaira, G. S.; Qin, J.; Garner, G. P.; Xiong, S.; Wan, L.; Ruiz, R.; Jaeger, H. M.; Nealey, P. F.; de Pablo, J. J. Evolutionary Optimization of Directed Self-Assembly of Triblock Copolymers on Chemically Patterned Substrates. *ACS Macro Lett.* **2014**, *3* (8), 747−752.
(57) Srinivasan, B.; Vo, T.; Zhang, Y.; Gang, O.; Kumar, S.; Venkatasubramanian, V. Designing DNA-Grafted Particles That Self-Assemble into Desired Crystalline Structures Using the Genetic Algorithm. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (46), 18431−18435.
(58) Fornleitner, J.; Lo Verso, F.; Kahl, G.; Likos, C. N. Genetic Algorithms Predict Formation of Exotic Ordered Configurations for Two-Component Dipolar Monolayers. *Soft Matter* **2008**, *4* (3), 480−484.
(59) Müller, K.; Osterman, N.; Babič, D.; Likos, C. N.; Dobnikar, J.; Nikoubashman, A. Pattern Formation and Coarse-Graining in Two-Dimensional Colloids Driven by Multiaxial Magnetic Fields. *Langmuir* **2014**, *30* (18), 5088−5096.
(60) Nikoubashman, A.; Likos, C. N. Self-Assembled Structures of Gaussian Nematic Particles. *J. Phys.: Condens. Matter* **2010**, *22* (10), 104107.
(61) Abraham, N. L.; Probert, M. I. J. A Periodic Genetic Algorithm with Real-Space Representation for Crystal Structure and Polymorph Prediction. *Phys. Rev. B* **2006**, *73* (22), 224104.
(62) Zhang, J.; Wang, C.-Z.; Ho, K.-M. Finding the Low-Energy Structures of Si[001] Symmetric Tilted Grain Boundaries with a Genetic Algorithm. *Phys. Rev. B* **2009**, *80* (17), 174102.
(63) Chua, A. L.-S.; Benedek, N. A.; Chen, L.; Finnis, M. W.; Sutton, A. P. A Genetic Algorithm for Predicting the Structures of Interfaces in Multicomponent Systems. *Nat. Mater.* **2010**, *9* (5), 418−422.
(64) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19* (14), 1639−1662.
(65) Wu, S. Q.; Ji, M.; Wang, C. Z.; Nguyen, M. C.; Zhao, X.; Umemoto, K.; Wentzcovitch, R. M.; Ho, K. M. An Adaptive Genetic Algorithm for Crystal Structure Prediction. *J. Phys.: Condens. Matter* **2013**, *26* (3), 035402.
(66) Curtis, F.; Li, X.; Rose, T.; Vázquez-Mayagoitia, Á.; Bhattacharya, S.; Ghiringhelli, L. M.; Marom, N. GAtor: A First-Principles Genetic Algorithm for Molecular Crystal Structure Prediction. *J. Chem. Theory Comput.* **2018**, *14* (4), 2246−2264.
(67) Spglib — Spglib v.2.0.0. https://spglib.github.io/spglib/ (accessed 2022−08−23).
(68) Li, X.; Curtis, F. S.; Rose, T.; Schober, C.; Vazquez-Mayagoitia, A.; Reuter, K.; Oberhofer, H.; Marom, N. Genarris: Random Generation of Molecular Crystal Structures and Fast Screening with a Harris Approximation. *J. Chem. Phys.* **2018**, *148* (24), 241701.
(69) Bier, I.; O'Connor, D.; Hsieh, Y.-T.; Wen, W.; Hiszpanski, A. M.; Han, T. Y.-J.; Marom, N. Crystal Structure Prediction of Energetic Materials and a Twisted Arene with Genarris and GAtor. *CrystEngComm* **2021**, *23* (35), 6023−6038.
(70) Li, Z.; Rosenmann, D.; Czaplewski, D. A.; Yang, X.; Gao, J. Strong Circular Dichroism in Chiral Plasmonic Metasurfaces Optimized by Micro-Genetic Algorithm. *Opt. Express* **2019**, *27* (20), 28313−28323.
(71) Redington, R. L.; Olson, W. B.; Cross, P. C. Studies of Hydrogen Peroxide: The Infrared Spectrum and the Internal Rotation Problem. *J. Chem. Phys.* **1962**, *36* (5), 1311−1326.
(72) Hunt, R. H.; Leacock, R. A.; Peters, C. W.; Hecht, K. T. Internal-Rotation in Hydrogen Peroxide: The Far-Infrared Spectrum and the Determination of the Hindering Potential. *J. Chem. Phys.* **1965**, *42* (6), 1931−1946.
(73) Cukier, R. I.; Seibold, S. A. Molecular Dynamics Simulations of Prostaglandin Endoperoxide H Synthase-1. Role of Water and the Mechanism of Compound I Formation from Hydrogen Peroxide. *J. Phys. Chem. B* **2002**, *106* (46), 12031−12044.
(74) Stillinger, F. H. Lattice Sums and Their Phase Diagram Implications for the Classical Lennard-Jones Model. *J. Chem. Phys.* **2001**, *115* (11), 5208−5212.
(75) Petsev, N. D.; Nikoubashman, A.; Latinwo, F. *Genetic Algorithm Crystal Optimization in a Chiral Tetramer Model System*; Princeton University, 2022. DOI: 10.34770/CHAG-5F03
(76) Mighell, A. D.; Karen, V. L. NIST Crystallographic Databases for Research and Analysis. *J. Res. Natl. Inst. Stand. Technol.* **1996**, *101* (3), 273−280.
(77) Mighell, A. D. Lattice Symmetry and Identification—The Fundamental Role of Reduced Cells in Materials Characterization. *J. Res. Natl. Inst. Stand. Technol.* **2001**, *106* (6), 983−995.
(78) Ising, E. Beitrag zur Theorie des Ferromagnetismus. *Z. Phys.* **1925**, *31*, 253−258.
(79) McCoy, B. M.; Wu, T. T. *The Two-Dimensional Ising Model*: 2nd ed.; Courier Corporation, 2014.
(80) Stillinger, F. H. Critical Point Confluence Phenomenon. *J. Phys. Chem. B* **2018**, *122* (13), 3441−3446.
(81) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117* (1), 1−19.
(82) Petsev, N. D.; Stillinger, F. H.; Debenedetti, P. G. *Chiral Tetramer Molecular Model Source Code*; Princeton University, 2021. DOI: 10.34770/ybs9-5t93