



Nanoscale

Accelerating Copolymer Inverse Design using Monte Carlo Tree Search

Journal:	<i>Nanoscale</i>
Manuscript ID	NR-ART-08-2020-006091.R2
Article Type:	Paper
Date Submitted by the Author:	12-Nov-2020
Complete List of Authors:	Patra, Tarak; Argonne National Laboratory, Center for Nanoscale Materials Loeffler, Troy; Argonne National laboratory, Center for Nanoscale Materials Sankaranarayanan, Subramanian K. R. S.; Argonne Natl Lab

SCHOLARONE™
Manuscripts

Accelerating Copolymer Inverse Design using Monte Carlo Tree Search

Tarak K Patra^{1*}, Troy D. Loeffler² and Subramanian K R S Sankaranarayanan^{2, 3*}

¹Department of Chemical Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu 600036, India

²Center for Nanoscale Materials, Argonne National Laboratory, Lemont, Illinois 60439, USA

³Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Chicago, Illinois 60607, USA

Abstract

There exists a broad class of sequencing problems in soft materials such as proteins and polymers that can be formulated as a heuristic search algorithm that involve decision making akin to a computer game. AI gaming algorithms such as Monte Carlo tree search (MCTS) gained prominence after their exemplary performance in the computer Go game and are decision trees aimed at identifying the path (moves) that should be taken by the policy to reach the final winning or optimal solution. Major challenges in inverse sequencing problems are that the materials search space is extremely vast and property evaluation for each sequence is computationally demanding. Reaching an optimal solution by minimizing the total number of evaluations in a given design cycle is therefore highly desirable. We demonstrate that one can adopt this approach for solving the sequencing problem by developing and growing a decision tree, where each node in the tree is a candidate sequence whose fitness is directly evaluated by molecular simulations. We interface MCTS with MD simulations and use a representative example of designing a copolymer compatibilizer, where the goal is to identify sequence specific copolymers that lead to zero interfacial energy between two immiscible homopolymers. We apply the MCTS algorithm to polymer chain lengths varying from 10-mer to 30-mer, wherein the overall search space varies from 2^9 (512) to 2^{29} (~0.5 billion). In each case, we identify a target sequence that leads to zero interfacial energy within a few hundred evaluations demonstrating the scalability and efficiency of MCTS in exploring practical materials design problems with exceedingly vast chemical/material search space. Our MCTS-MD framework can be easily extended to several other polymer and protein inverse design problems, in particular, for cases where sequence-property data is either unavailable and/or is resource intensive.

Key words: Materials Design, Machine Learning, Monte Carlo Tree Search, and Molecular Dynamics Simulation

*E-mail: tpatra@iitm.ac.in (TKP), skrssank@uic.edu (SKRSS)

Introduction

There exists a broad class of soft-materials such as proteins and polymers where the arrangement of moieties *i.e.* the sequence plays a critical role in determining their functionality. For instance, the activities and functionalities of DNA and other biomolecules are determined by the exact sequence of amino acids and other chemical moieties in their back bones.¹⁻³ As an example, the arrangement of amino acid sequence in viruses plays a key role in determining their mutations and hence the effectiveness of the drugs or vaccine used to treat them. Likewise, several recent studies indicate that the sequence specificity of the constituents chemical moieties of a copolymer can lead to more efficient materials – their thermodynamic properties such as miscibility and surface tension as well as structure/morphology is strongly influenced by the sequence in oligomers. It is therefore not surprising that a lot of effort has focused on controlling the sequence specificity in polymers, proteins and other biomolecules.

On the experimental front, progress in synthetic chemistry has enabled us to exercise an unprecedented control over sequences in copolymers - such precision polymers remain an area of major focus in current fundamental and applied polymer research.⁴⁻⁶ Copolymers are a special class of polymers that comprise of more than one type of chemical species; and shows a reach phase behaviour⁷⁻⁹ and tunability in its thermophysical properties.¹⁰⁻¹⁵ These copolymers are usually characterized by their mean block length and mass fraction. One of the area where sequence specificity is found to play an import role is the use of copolymer as interfacial compatibilizers.^{16,17} Copolymer compatibilizers are commonly employed to improve the thermodynamic stability of polymer interfaces, and they therefore have wide applicability in emulsions and composite materials.^{11,18}

A major challenge in the design of sequence specific polymers lies in the vast combinatorial space that precludes efficient exploration. For instance, a polymer chain with n number of possible monomers and m type of monomers will have n^m possible combinations that can be likely explored. Even for a binary polymer *i.e.* 2 types with a chain length of ~ 30 units, the total combinations possible (accounting for double counting) are 2^{29} which is close to 0.5 billion. Given such an enormous sequence space that needs to be explored, it is highly desirable to minimize the number of trials needed to arrive at a sequence that corresponds to a desired target property. Fortunately, the emergence of artificial intelligence (AI) positions us uniquely to solve this seemingly intractable inverse design problem.

AI and machine learning (ML) has been increasingly interfaced with molecular simulations to solve the inverse problem and accelerate materials discovery/design. Molecular simulations such as molecular dynamics (MD) are powerful techniques to evaluate the

sequence-property relationships. Typically, MD simulations can sample the configurational & property space and create adequately large structure-property training datasets of a materials. On the other hand, ML methods can very efficiently screen this extensive dataset and identify sequences or configurations that correspond to desired optimal material properties. Such inverse problems have been traditionally addressed using evolutionary methods such as genetic algorithms (GA) or Bayesian optimization (BO) – these optimizers are combined with MD simulations to identify target properties for a wide range of materials from inorganic to semiconductor to polymers.^{17,19–23} In the last few years, such combination of MD and ML have been successfully deployed to explore the vast configurational space of materials. Their widespread application to problems of practical interest, however, requires addressing two bottlenecks. First, the GA and BO methods exhibit poor scalability as the design space increases.²⁴ Typically, the search space of most practical sequencing problems exceeds several millions and higher. Second, they find difficulties in surmounting suboptimal solutions and tends to slow down near the optimal points. Third, each property evaluation in many design problems is computationally intensive (trajectories over several tens of nanoseconds and more), which precludes high-throughput exploration.

Within a typical MD-ML materials design framework, one often requires several thousands to millions of direct evaluations or computations of materials properties. In the context of soft materials, this poses a major limitation, for instance, when the MD calculations for each sequence are computational very expensive such as, for instance, computing the thermophysical properties of polymeric materials. We note that the relaxation in polymeric materials is inherently slow and requires significantly long MD simulations to calculate their equilibrium properties. A key challenge in accelerating computer aided molecular-scale polymer design and address the sequence problems in materials design is to significantly reduce the number of direct computational evaluation of materials property that are required to identify an optimal candidate corresponding to the target property.

The advent of big data analytics and powerful supercomputers have brought AI and ML techniques that can address the above challenges in materials design. In this front, Monte Carlo tress search (MCTS) has emerged as a powerful global optimization method that has found wide-spread applications in computer games such as Alpha Go, games such as Bridge, Poker and many other video games.^{25,26} MCTS is a probabilistic and heuristic search algorithm that integrates a tree search algorithm with machine learning principles of reinforcement learning. MCTS is a decision tree-based approach that builds a shallow tree of nodes where each node represents a point in the search space and downstream pathways are generated by a rollout

procedure. The algorithm simultaneously explores potentially better pathways to reach the optimal point in a search space and exploits a single pathway that has the greatest estimate value of the search function. This combination of exploration vs. exploitation and an appropriate trade-off mechanism between them are found to be the most efficient strategy of identifying optimal point for a given function. An advantage of the MCTS is that if the search gets trapped in a metastable or suboptimal point, it can quickly find another pathway by growing other branches of the tree utilizing the trade-off mechanism between exploration and exploitation. Recently, MCTS has been successfully adopted for material science problems such as predicting silicon-germanium alloy structure with optimal thermal conductance,^{24,27} and discovering new synthetic routes of making organic molecules,²⁸ optimal atoms segregation at grain boundary,²⁹ predicting organic molecules with optimal partition coefficient and other properties,³⁰ and enhancing biomolecular sampling.³¹

Here, we draw inspiration from the recent success of AI algorithms such as MCTS in computer games and aim to develop a design algorithm for sequence problems that are fast (time to solution) as well as highly scalable. We focus on a representative albeit complex polymer inverse design problem, viz., design of sequence of copolymer molecules that corresponds to a user-desired property. Our goal is to design the sequence of the compatibilizer that minimizes the interfacial tension between immiscible polymers. Block copolymers and random copolymers have long been used as compatibilizers that reduce interfacial tension between immiscible polymers and improve the stability of the composite materials.^{16,32} These copolymers manipulate nanoscale domain structure and interaction to enhance the stability and mechanical properties of composite materials. Recently, evolutionary search based on MD simulations (MD-GA) have identified sequence specific co-polymers that outperform block and random copolymers.¹⁷ However, an optimal solution (a sequence specific copolymer) for a 20 bead polymer chain via evolutionary search required several thousands of MD simulations within the MD-GA framework. In practice, the polymer chains can often involve several tens of monomers to several hundreds necessitating algorithms that are efficient and scalable. In our design workflow, we interface MCTS with MD simulations that is used for evaluating the objective function for any specific sequence. Our MD simulation based Monte Carlo tree search (MD-MCTS) workflow rapidly identifies optimal sequence of copolymers corresponding to our desired interfacial tension. We demonstrate the scalability of our workflow by simulating chain lengths from 10 to 30 monomers – for each case the search required only a few hundred evaluations despite the search space extending from 512 to 0.5

billion, respectively. Our work demonstrates the success of AI in efficient and faster materials search and is applicable for a broad class of sequence related materials design problems.

Model and Methodology

Molecular Dynamics of Polymers: We use a generic coarse-grained model to represent two homopolymers A and B that are immiscible. The compatibilizer is a copolymer consist with both the A and B type moieties. Within this model system, two adjacent coarse-grained monomers of a polymer is connected by the Finitely Extensible Nonlinear Elastic (FENE) potential of the form

$$E = -\frac{1}{2}KR_0^2\left[1 - \left(\frac{r}{R_0}\right)^2\right].$$

Here, $k = 30\epsilon/\sigma^2$ and $R_0 = 1.5\sigma$. Any two monomers in the system is interacted via the Lennard-Jones (LJ) potential of the form

$$V(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^6\right].$$

The ϵ_{ij} is the interaction energy between any two monomers i and j . The size of all the monomers are σ . The LJ interaction is truncated at a cut-off distance $r_c = 2.5\sigma$ to represent attractive interaction among the monomers of homopolymers viz., A-A and B-B interactions. The immiscibility between homopolymer A and B is modelled by pure repulsion between A and B moieties. This is achieved by choosing $r_c = 2^{1/6}\sigma$ for A-B interaction. The orthogonal simulations box consist of a total of 693 homopolymer A and 693 homopolymer B that form two interfaces as shown in Figure 1a. The compatibilizer chains are placed at both the interfaces. The interfacial area between the two homopolymers ($36\sigma \times 36\sigma$) is kept constant during the simulations. Also the compatibilizer concentration at each interface which is defined as the compatibilizer monomers per unit area of an interface, is kept constant. Three case studies are conducted each for varying compatibilizer chain length. A total of 414, 207 and 138 compatibilizer copolymer chains of length $N=10, 20$ and 30 , respectively, are placed at both the interfaces. This lead to a compatibilizer density of $1.59/\sigma^2$. All the systems consist of 36000 CG beads. The system is periodic in all three direction. All the simulations employ the Verlet time integration scheme³³ with a time step of 0.005τ , where the unit of time is $\tau = \sigma\sqrt{m/\epsilon}$. The Nose-Hover thermostat and barostat³⁴ are employed to keep the temperature and pressure constant during the simulations.

Property Evaluation: All the MD calculations are conducted at a reduce temperature $T=1$ and zero pressure in the direction normal to the interface. During a MD summation, a system is initially equilibrated for 2×10^6 MD steps, followed by a production run of another 2×10^6 steps. During the production run, pressure tensor data are collected, and the surface tension of a system is calculated as

$$\gamma_{12} = \left\langle \frac{L_z}{2} \left(P_{zz} - \frac{1}{2}(P_{xx} + P_{yy}) \right) \right\rangle.$$

Here z is the direction normal to the interface, the in-equilibrium box length along z is represented by L_z . The P_{xx} , P_{yy} and P_{zz} are the pressure components along three directions. All the MD simulations are conducted using LAMMPS molecular dynamics simulation package.³⁵

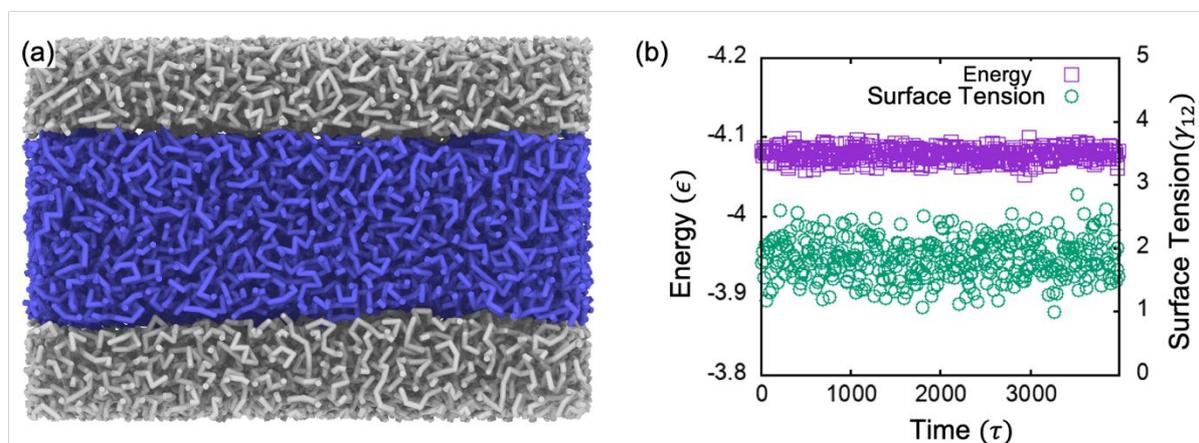


Figure 1: A polymer blend: (a) MD snapshot of two immiscible homopolymers forming interfaces in a MD simulation box. (b) energy and surface tension of the system is shown as a function of time during the production run. The energy, time and surface tension values are in LJ units.

Material Systems: We first model a base material where two immiscible homopolymers of type A and B form interfaces, as shown by the MD snapshot in Figure 1a. A representative MD simulation depicting in-equilibrium energy and surface tension as a function of time is shown in Figure 1b. The system attains a steady state within the equilibrium run and all the properties are calculated by time averaging of data collected in these equilibrium region of the trajectory. The time average energy and surface tension of the system are -4.09ϵ and $1.8\epsilon/\sigma^2$, respectively. Here, ϵ and σ are the unit of energy and length, respectively. Next, the copolymer chains of both A and B type moieties, at the interfaces, are simulated. We subsequently interface the MD simulations with MCTS to explore the sequence of A and B type moieties in a given copolymer or compatibilizer that reduces the surface tension of the system. Three compatibilizer chains of length $N = 10, 20$ and 30 are considered and are shown in Figure 2. The total number of possible candidate structures for a binary chain is $C = 2^N/2$. The

denominator 2 is to avoid the double counting of configurations - a sequence and its reverse sequence are identical in this context. Therefore, the total number of candidate structures or sequences in the search space are 512, 524288 and 536870912 for $N=10$, 20 and 30, respectively. We seek to identify the sequence of moieties A and B that lead to lowest surface tension of the system for all the three cases by combining MD and MCTS.

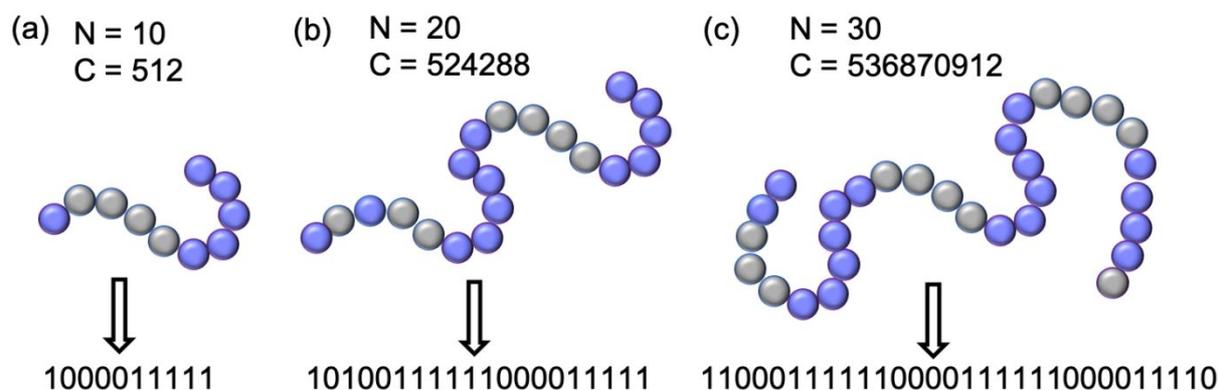


Figure 2: Binary Mapping of copolymers. Copolymer of length $N=10, 20$ and 30 are shown in (a), (b) and (c), respectively. The grey and blue beads represent monomer of type A and B, respectively. The arrows point to the one dimensional binary strings where 0 and 1 correspond to monomer A and B, respectively. The binary string length is same as the copolymer chain length N . Here, the C represent total number of sequences possible for a given copolymer of size N .

Monte Carlo Tree Search For Co-polymer Design: The MD-MCTS workflow for exploring the sequence search space is shown schematically in Fig. 3. The objective is to minimize the surface tension of the system in the presence of compatibilizer chains – the surface tension can be written as $\gamma_{12} = \gamma_{12}(x)$. Here, $x \in \{0, 1\}^N$ represent a sequence of 0 and 1 of size N and 0 and 1 correspond to moieties A and B, respectively. We conduct three MD-MCTS calculations each for a specific value of N . For a given N , the MD-MCTS begins with randomly generating a sequence of 0 and 1 of size N . This candidate serves as the root node of the search tree. A

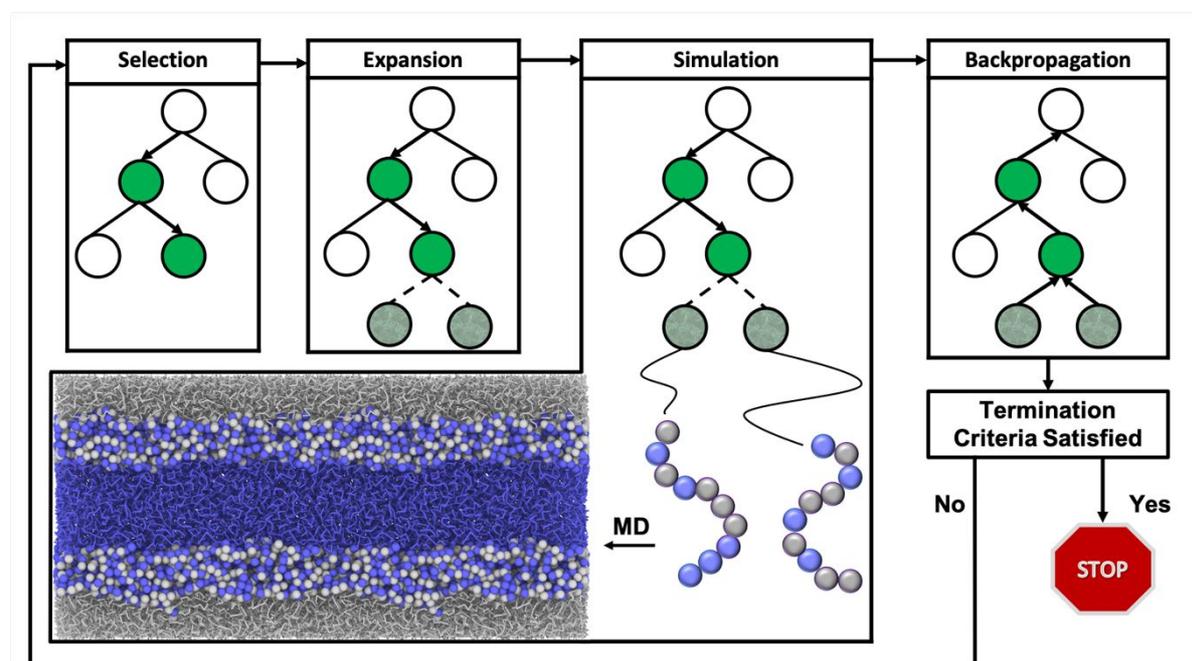


Figure 3: MD-MCTS design scheme for copolymer. It comprise of four steps – selection, expansion simulation and backpropagation that are sequentially conducted as shown by the arrows in a given iteration. In the simulation step, MD calculations of a set of candidate structures are conducted parallelly. In the MD snapshot, homopolymers are shown as lines; and copolymers are shown as beads. The termination criteria is chosen to be $\gamma_{12} \approx 0.0$.

search tree is built in an incremental and iterative way for searching optimal sequence of a copolymer as shown in Figure 3. Each node of the tree represents a specific sequence of the copolymer. Here the sequence is represented by a binary number of digit N for a copolymer of chain length N . Once the termination criteria is reached (set to be $\gamma_{12} \approx 0$), the search (growth of the tree) is stopped and the best performing candidate is returned. In each iteration, four steps – selection, expansion, simulations and back-propagation are carried out. A child node is selected during the selection process based on the upper confidence bound (UCB) score.³⁶ The UCB of a node is defined as $ucb_i = \frac{\sum_{k=1}^{N_i} \gamma_k}{v_i} + C \sqrt{\frac{2 \ln v_p}{v_i}}$. Here, γ_k is the surface tension of the k -th playout performed by this node and all of its downstream child nodes, and v_i is the visit/playout count of the node, v_p is the visit/playout count of the parent of this node and C is a constant for balancing the exploration side (the right hand side of the plus sign) and the exploitation (the left hand side of the plus sign). All variables in this equation are aggregates of a node and its child nodes with exception of the exploration constant. The exploration constant, C , is of course a hyper parameter that is chosen by the user. However this is a highly critical choice given that the efficiency of the algorithm will be determined by this. If the value of C is too small this will cause the exploitation side of the equation to be the dominant term and as a result nodes will only be selected according to which node has the best score at this moment. This will of course cause the algorithm to quickly flow to the nearest local minima and get trapped. If the value is too large the selection process will wander aimlessly. Note that the node with the highest UCB score is always selected for a given step. The MCTS algorithm doesn't directly pick to exploit or explore, it instead picks according to the best combined total. For a node to have a high UCB score, the node must have a good combination of both terms. A good choice of the exploration constant is critical as this determines how these two terms combine to give a UCB score.

The primary benefit of the MCTS formalism is that when properly tuned the algorithm will make a weighted choice to continue down a path that appears to have a good solution hidden behind it or to cut its losses and look somewhere else when it appears to have exhausted its search space down a given path. As such the choice of the exploration constant is very critical for optimal performance as it will search a path well enough to potentially find a solution, but not waste valuable simulation time by over-sampling a given path. Instead of trying to manually tune the parameter, the value of C was controlled adaptively at each node according to the formula $C = \frac{\sqrt{2}J}{4}(z_{max} - z_{min})$. Here, J is the meta parameter which is set to be

one and it increases whenever the algorithm reach a “dead end” node to allow more exploration. At a dead-end node, the number of possible structures narrows to one. This happens when the numbers of $k - 1$ candidate structures reach the limit. Here, the J is updated as $J \leftarrow J + \max \left\{ \frac{T-t}{T}, 0.1 \right\}$, where T is the total number of candidates to be evaluated, and t is the number of candidates for which the surface tension is already evaluated. Whenever a new node is added, it is by default selected for one simulation cycle as part of the initialization process. Next, we perform the expansion of the tree by adding child nodes to the selected node. A new child node is created by randomly flipping a digit of the parent node. In the simulation step, a playout is performed from each of the added children. We roll out 10 structures randomly during a playout from a child node. In these playouts the structure contained within this node is changed by randomly flipping several polymer groups. Initially the entire chain from the initial structure is allowed to flip randomly, but for nodes that are deeping in the tree (IE 2nd, 3rd, or 4th generation nodes) fewer polymer groups are allowed to flip. This ensures a convergent behaviour where less and less of the chain is modified as the algorithm picks a path in the node tree to travel down. The scaling was set to be 100%, 60%, 30%, 20%, 10%, 5% for the 0th, 1st, 2nd, 3rd, etc. generation node respectively in the tree. The final level was always set such that only 1 polymer unit was flipped. All 10 playout structures are evaluated *via* long-time scale MD simulations.

Finally, in the back propagation step, the visit count of each ancestor node of i is incremented by one and the cumulative value is also updated to keep consistency. Note that the concept of back propagation (i.e. child nodes feeding their information back up the tree) is a key feature of this algorithm. Owing to this, a parent also shares the reward that is discovered by its child node. As such, the reward for every node, the branch starting from the child who discovered it up to the head node is updated. This makes it possible for the algorithm to either choose to simply flip a few polymer units by selecting the deepest child node or return back up the tree to one of the higher parent nodes in the branch, where it has more freedom in how many flips one can perform. This effectively means the algorithm is allowed to pick how many flips it wishes to try by selecting a node of a given depth. This balance allows it to both make fine-tuned adjustments or larger adjustments as needed. In the absence of this, either the algorithm never converged because it was changing too much all the time or it would become trapped because it was not making a large enough change to the polymer.

The initial layer of nodes effectively is given a random sequence of polymer chains and the child nodes are refined based on the reward. MCTS then picks between either large changes

or small ones depending on finding sufficient reward down a particular branch. If a child node is failing to find better rewards than its parent, the UCB will start favoring the parent node instead of the child node. This will result in the algorithm making larger adjustments instead of smaller ones. (i.e. single polymer unit flips are failing to result in further minimize the surface tension suggesting larger number of flips are required). At any given point, any node in the tree regardless of depth, can be chosen if its UCB score becomes the largest. As such, it is possible to sample any point in the total statistical space. Finally, we point out that the depth scaling rates are an arbitrary hyperparameter, but in this case were chosen to ensure a smooth transition between all the polymer units being randomly assigned and only 1 being flipped. We find this is obtained by a tree depth of 6 nodes which strikes a good balance between speed and accuracy.

Results:

We first assess the performance of our MCTS-MD workflow as shown in Figure 4. The lowest surface tension as a function of total number of candidate evaluated during the MCTS iterations is shown in Figure 4a for the three different polymer chain lengths. We find that the MD-MCTS is able to identify optimal sequences that bring down the surface tension to near zero for each of the three cases. An optimal sequence for $\gamma_{12} \approx 0.0$ for each case is shown in Figure 4b. It is interesting to note that all the optimal sequences are non-periodic and non-intuitive. These machine learned sequences combine small and long segments of blocks. These specific arrangement of blocks in the compatibilizer structure lead to more interfacial crossings at the interface than regular di-block copolymers or other periodic copolymers of large block length. The ML identified polymers outperform di-block copolymers due to their ability to form large number of interfacial crossings at the interface. MCTS demonstrates exceptional scalability

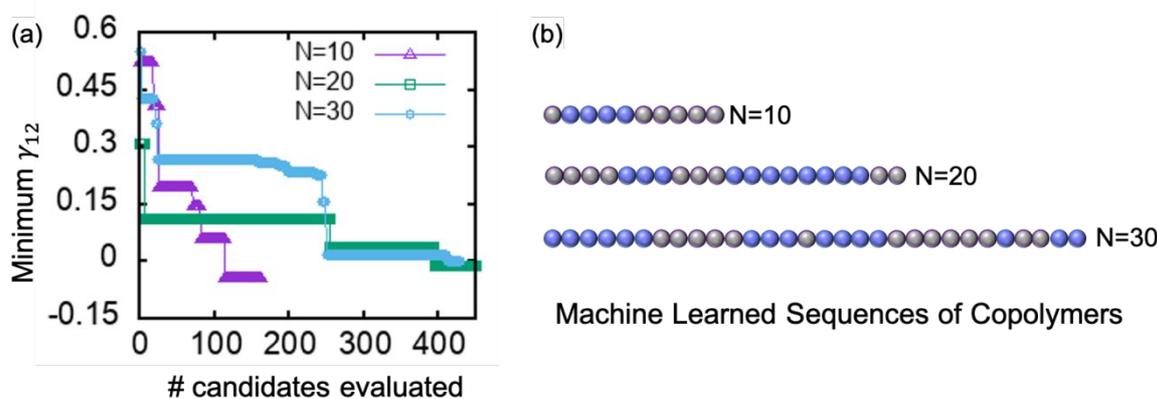


Figure 4: Performance and prediction of MD-MCTS. (a) The lowest surface tension achieved during the MD-MCTS run is shown as a function of total number of candidate materials directly evaluated via MD simulations for all three cases. The optimal sequences for copolymers for all three cases are depicted in (b).

and is able to achieve zero interfacial energy irrespective of the size or chain length of the polymer compatibilizer

We assess the scalability of the MD-MCTS algorithm by plotting the total number of candidate structures evaluated during a given search cycle as a function of the size of the compatibilizer (Figure 5a). This is especially important considering the long timescale MD simulations that are necessary to perform each evaluation for a specific sequence. MCTS performs direct evaluation of materials properties of 114, 388 and 415 candidates to achieve an optimal sequence that correspond to our target *i.e.* zero interfacial energy of the system for copolymers of length 10, 20 and 30, respectively. We only note a marginal increase in the number of MD simulations with an increase in the system size. This is incredible considering that the design space *i.e.* the total number of candidates, C increase from 1024 to 1 billion when the chain lengths increase from 10 to 30. MCTS is thus able to attain an optimal solution by screening lower percentage of candidate *viz.*, 22%, 0.07 % and 7.7×10^{-5} % of total possible structures for chain length 10, 20 and 30, respectively. Figure 5b depicts this ratio C_s/C as a function of

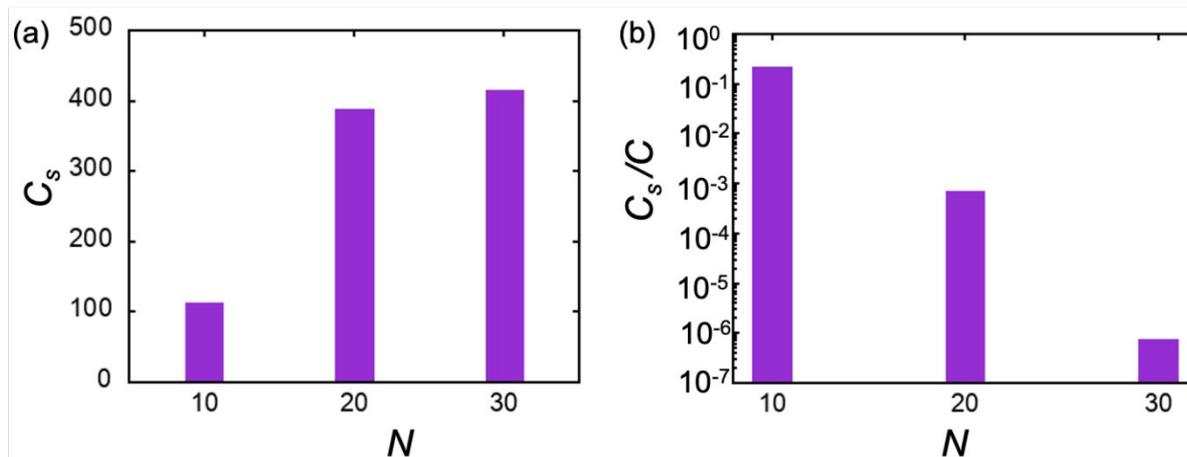


Figure 5: MD-MCTS complexity. (a) The total number of candidate structures screened (C_s) during a MCTS cycle is shown as a function of number of monomer in a copolymer chain. (b) The fraction of candidate structures (C_s/C) screened during a design cycle is plotted as a function of polymer chain length.

N , clearly indicating the exponentially lower fraction of candidates required to be screened during the design cycle as polymer chain length increases. This strongly suggests that the MD-MCTS design scheme is scalable to extremely large system sizes, which has hitherto posed a challenge to evolutionary search strategies.

Mean block length of a copolymer has long been perceived as an important descriptor of a copolymer's properties. Therefore, we closely analyse the correlation between any given sequence *i.e.* the relative statistics of type-1 and type-0 moieties present in a copolymer chain and the computed surface tension of the system. Surface tension vs. mean block length for each of the three systems ($N=10, 20$ and 30) is shown in Fig. 6. Here, the mean block length is calculated as the arithmetic mean of the size of all the blocks of 0's and 1's presence in a copolymer. The MCTS identifies multiple mean block lengths that correspond to highest performance compatibilizer ($\gamma_{12} \approx 0$). For example, there are copolymers of size $N=10$ with mean block length $b_l = 1.8, 2.6$ and 3.3 that yield $\gamma_{12} \approx 0$ (Fig. 6a). Similar observation can be made for $N=20$ and $N=30$ from Figure 6b and 6c, respectively. We also find that sequences with the same mean block length as the optimal sequence can exhibit a wide range of interfacial energies. The interfacial energy varies from 0 to $1.2\epsilon/\sigma^2$, approximately, for a mean block length of 3.3 for $N=10$. A similar variation of surface tension for a given mean block is observed for $N=20$ and 30 as evident in Figure 6b, and 6c, respectively. Thus, the mean block length alone is a poor predictor of a compatibilizer performance.

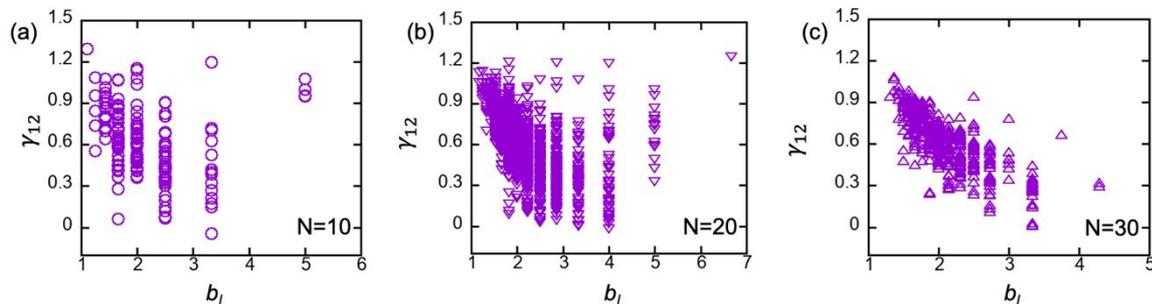


Figure 6: Variation of interfacial energy is shown as a function of the mean block length for all the candidate structures screened in this study for $N = 10, 20$ and 30 in (a), (b) and (c), respectively. Both the surface tension and mean block length are in LJ unit.

To further understand the uniqueness of the optimal polymer sequence, we study the interrelationship between mean block length, monomer mole fraction and the interfacial energy of the system. Here, the mole fraction is defined as the ratio of the number of monomer of one particular moiety *viz.* type-1 moiety to the total number of monomers in a chain. Figure 7a, b and c show the variation of interfacial energy for $N=10, 20$ and 30 , respectively, as a function of mean block length (l_b) and mole fraction of type-1 (q). The deep blue contours in Figure 7 corresponds to lowest surface tension of the system. For $N=10$, we observe $\gamma_{12} \approx 0.0$ for $q \approx 0.5$ and $l_b = 3.3$. As N increases, there are larger patches of isolated blue contours that appear in the q - l_b surface, indicating greater number of global optimal points in the q - l_b surface.

As the chain length further increases, it is possible to achieve many more sequences that can lead to zero interfacial energy of the system. Often times, the optimal regions are separated by large regions of suboptimal points as evident from Figure 7b and 7c. Many of the local optimization strategies and even global ones struggle to navigate around these sub-optimal points. Thus, the complexity as well as the degeneracy of this configurational space increases for longer chain length and poses challenges. The MCTS algorithm is however able to effectively navigate around these sub-optimal regions by growing other branches of the tree effectively utilizing the trade-off mechanism between exploration and exploitation. MCTS thus simultaneously explores potentially better pathways to reach the optimal point in a search space and exploits pathways that have the greatest estimate value of the search function. This combination of exploration vs. exploitation and an appropriate trade-off mechanism between them, represents a powerful strategy of identifying optimal point for a given function.

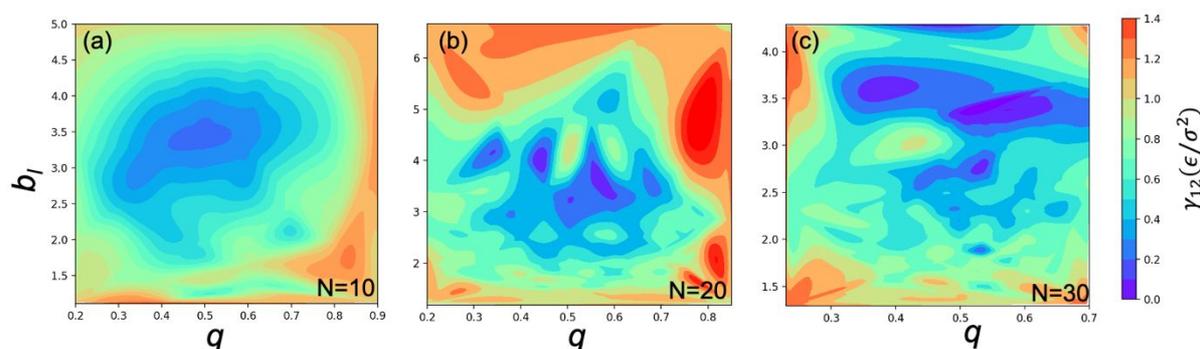


Figure 7: Heat map of interfacial energy. The interfacial energy γ_{12} is shown as a function of mole fraction of A type monomer, q , and mean block length, b_l . The surface tension and block length are in LJ unit.

Conclusions

In summary, sequence control in soft matter systems such as polymers and proteins has been a longstanding goal and is highly desirable for a wide variety of applications. In particular, the emergence of sequence control polymers provides tremendous opportunities for material design. By controlling the sequence of a copolymer, one can reliably tune their functionality over a wide range which can be mapped on to their chemical information. The vast combinatorial search space that is required to be explored for these sequence problems pose a major challenge. To overcome this, we interface a gaming AI algorithm viz., MCTS with an MD simulator to tackle the complexity of copolymer design. Our MD-MCTS design algorithm is employed to identify optimal copolymers sequences which can be used as compatibilizer to improve thermodynamics stability of two immiscible homopolymer blends. Unlike other molecular inverse design strategies *i.e.* genetic algorithm (GA) or Bayesian optimization (BO)

where design time grows rapidly with the system size, MCTS interfaced with MD appears to require relatively much smaller number of candidate evaluations in any given design cycle. We show that one can engineer the sequence of chemical moieties that will nullify the interfacial tension between immiscible polymers irrespective of the size of compatibilizer polymer chain. Our work also elucidates the correlation between interfacial energy and the sequence statistics of copolymer compatibilizer molecules and illustrates the complexities associated with sequence control over larger polymer chain lengths. Finally, we show that MCTS is highly scalable and efficiently able to navigate large design spaces typical of most practical sequence control related design problem. More broadly, the work provides new strategy that can be used for sequence control and inverse design of copolymers for materials applications.

Acknowledgement

The use of the Center for Nanoscale Materials, an Office of Science user facility, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. This research used resources of the National Energy Research Scientific Computing Center, which was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program of the Argonne Leadership Computing Facility at the Argonne National Laboratory, which was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-06CH11357. The authors would like to acknowledge the support from the Argonne LDRD-2017-012-N0 project, UIC faculty start-up fund and IIT Madras faculty stat-up fund. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences Data, Artificial Intelligence and Machine Learning at DOE Scientific User Facilities program under Award Number 34532.

Author contributions

T.P., T.L. and S.K.R.S. conceived and designed the project. T.P. and T.L. contributed equally. T.P., T.L. and S.K.R.S. developed the inverse design framework using MCTS. T.L. developed the MCTS workflow while T.P. carried out the MCTS optimization. T.P. carried out the MD simulations of polymers. All authors interpreted the data and performed analysis. T.P. and

S.K.R.S. wrote the manuscript. All the authors commented and contributed to the preparation of the manuscript.

Competing interests

The authors declare no competing financial or non-financial interests.

Reference:

- (1) Gruenbaum, Y.; Naveh-Many, T.; Cedar, H.; Razin, A. Sequence Specificity of Methylation in Higher Plant DNA. *Nature* **1981**, *292* (5826), 860–862. <https://doi.org/10.1038/292860a0>.
- (2) Mitchell, P. J.; Tjian, R. Transcriptional Regulation in Mammalian Cells by Sequence-Specific DNA Binding Proteins. *Science* **1989**, *245* (4916), 371–378. <https://doi.org/10.1126/science.2667136>.
- (3) Lemmon, M. A.; Flanagan, J. M.; Treutlein, H. R.; Zhang, J.; Engelman, D. M. Sequence Specificity in the Dimerization of Transmembrane α -Helices. *Biochemistry* **1992**, *31* (51), 12719–12725. <https://doi.org/10.1021/bi00166a002>.
- (4) Lutz, J.-F. Defining the Field of Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **2017**, *38* (24), 1700582. <https://doi.org/10.1002/marc.201700582>.
- (5) Neve, J. D.; Haven, J. J.; Maes, L.; Junkers, T. Sequence-Definition from Controlled Polymerization: The next Generation of Materials. *Polym. Chem.* **2018**, *9* (38), 4692–4705. <https://doi.org/10.1039/C8PY01190G>.
- (6) Perry, S. L.; Sing, C. E. 100th Anniversary of Macromolecular Science Viewpoint: Opportunities in the Physics of Sequence-Defined Polymers. *ACS Macro Lett.* **2020**, *9* (2), 216–225. <https://doi.org/10.1021/acsmacrolett.0c00002>.
- (7) Bates, F. S. Polymer-Polymer Phase Behavior. *Science* **1991**, *251* (4996), 898–905. <https://doi.org/10.1126/science.251.4996.898>.
- (8) Chremos, A.; Nikoubashman, A.; Panagiotopoulos, A. Z. Flory-Huggins Parameter χ , from Binary Mixtures of Lennard-Jones Particles to Block Copolymer Melts. *J. Chem. Phys.* **2014**, *140* (5), 054909. <https://doi.org/10.1063/1.4863331>.
- (9) Bates, F. S.; Fredrickson, G. H. Block Copolymer Thermodynamics: Theory and Experiment. *Annu. Rev. Phys. Chem.* **1990**, *41* (1), 525–557. <https://doi.org/10.1146/annurev.pc.41.100190.002521>.
- (10) Hamley, I. W. Ordering in Thin Films of Block Copolymers: Fundamentals to Potential Applications. *Prog. Polym. Sci.* **2009**, *34* (11), 1161–1210. <https://doi.org/10.1016/j.progpolymsci.2009.06.003>.
- (11) Cigana, P.; Favis, B. D.; Jerome, R. Diblock Copolymers as Emulsifying Agents in Polymer Blends: Influence of Molecular Weight, Architecture, and Chemical Composition. *J. Polym. Sci. Part B Polym. Phys.* **1996**, *34* (9), 1691–1700. [https://doi.org/10.1002/\(SICI\)1099-0488\(19960715\)34:9<1691::AID-POLB18>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1099-0488(19960715)34:9<1691::AID-POLB18>3.0.CO;2-2).
- (12) Meijer, H. E. H.; Lemstra, P. J.; Elemans, P. H. M. Structured Polymer Blends. *Makromol. Chem. Macromol. Symp.* **1988**, *16* (1), 113–135. <https://doi.org/10.1002/masy.19880160109>.
- (13) Sundararaj, U.; Macosko, C. W. Drop Breakup and Coalescence in Polymer Blends: The Effects of Concentration and Compatibilization. *Macromolecules* **1995**, *28* (8), 2647–2657. <https://doi.org/10.1021/ma00112a009>.

- (14) Khokhlov, A. R.; Khalatur, P. G. Conformation-Dependent Sequence Design (Engineering) of AB Copolymers. *Phys. Rev. Lett.* **1999**, *82* (17), 3456–3459. <https://doi.org/10.1103/PhysRevLett.82.3456>.
- (15) Sing, C. E.; Zwanikken, J. W.; Olvera de la Cruz, M. Electrostatic Control of Block Copolymer Morphology. *Nat. Mater.* **2014**, *13* (7), 694–698. <https://doi.org/10.1038/nmat4001>.
- (16) Lyatskaya, Y.; Gersappe, D.; Gross, N. A.; Balazs, A. C. Designing Compatibilizers To Reduce Interfacial Tension in Polymer Blends. *J. Phys. Chem.* **1996**, *100* (5), 1449–1458. <https://doi.org/10.1021/jp952422e>.
- (17) Meenakshisundaram, V.; Hung, J.-H.; Patra, T. K.; Simmons, D. S. Designing Sequence-Specific Copolymer Compatibilizers Using a Molecular-Dynamics-Simulation-Based Genetic Algorithm. *Macromolecules* **2017**, *50* (3), 1155–1166. <https://doi.org/10.1021/acs.macromol.6b01747>.
- (18) Macosko, C. W.; Guégan, P.; Khandpur, A. K.; Nakayama, A.; Marechal, P.; Inoue, T. Compatibilizers for Melt Blending: Premade Block Copolymers. *Macromolecules* **1996**, *29* (17), 5590–5598. <https://doi.org/10.1021/ma9602482>.
- (19) 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. <https://doi.org/10.1021/acscombsci.6b00136>.
- (20) Patra, T. K.; Zhang, F.; Schulman, D. S.; Chan, H.; Cherukara, M. J.; Terrones, M.; Das, S.; Narayanan, B.; Sankaranarayanan, S. K. R. S. Defect Dynamics in 2-D MoS₂ Probed by Using Machine Learning, Atomistic Simulations, and High-Resolution Microscopy. *ACS Nano* **2018**, *12* (8), 8006–8016. <https://doi.org/10.1021/acsnano.8b02844>.
- (21) Lookman, T.; Balachandran, P. V.; Xue, D.; Yuan, R. Active Learning in Materials Science with Emphasis on Adaptive Sampling Using Uncertainties for Targeted Design. *Npj Comput. Mater.* **2019**, *5* (1), 1–17. <https://doi.org/10.1038/s41524-019-0153-8>.
- (22) Jackson, N. E.; Webb, M. A.; de Pablo, J. J. Recent Advances in Machine Learning towards Multiscale Soft Materials Design. *Curr. Opin. Chem. Eng.* **2019**, *23*, 106–114. <https://doi.org/10.1016/j.coche.2019.03.005>.
- (23) Schmidt, J.; Marques, M. R. G.; Botti, S.; Marques, M. A. L. Recent Advances and Applications of Machine Learning in Solid-State Materials Science. *Npj Comput. Mater.* **2019**, *5* (1), 1–36. <https://doi.org/10.1038/s41524-019-0221-0>.
- (24) M. Dieb, T.; Ju, S.; Yoshizoe, K.; Hou, Z.; Shiomi, J.; Tsuda, K. MDTS: Automatic Complex Materials Design Using Monte Carlo Tree Search. *Sci. Technol. Adv. Mater.* **2017**, *18* (1), 498–503. <https://doi.org/10.1080/14686996.2017.1344083>.
- (25) Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; Dieleman, S.; Grewe, D.; Nham, J.; Kalchbrenner, N.; Sutskever, I.; Lillicrap, T.; Leach, M.; Kavukcuoglu, K.; Graepel, T.; Hassabis, D. Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature* **2016**, *529* (7587), 484–489. <https://doi.org/10.1038/nature16961>.
- (26) Browne, C. B.; Powley, E.; Whitehouse, D.; Lucas, S. M.; Cowling, P. I.; Rohlfshagen, P.; Tavener, S.; Perez, D.; Samothrakis, S.; Colton, S. A Survey of Monte Carlo Tree Search Methods. *IEEE Trans. Comput. Intell. AI Games* **2012**, *4* (1), 1–43. <https://doi.org/10.1109/TCIAIG.2012.2186810>.
- (27) Dieb, T. M.; Ju, S.; Shiomi, J.; Tsuda, K. Monte Carlo Tree Search for Materials Design and Discovery. *MRS Commun.* **2019**, *9* (2), 532–536. <https://doi.org/10.1557/mrc.2019.40>.

- (28) Segler, M. H. S.; Preuss, M.; Waller, M. P. Planning Chemical Syntheses with Deep Neural Networks and Symbolic AI. *Nature* **2018**, *555* (7698), 604–610. <https://doi.org/10.1038/nature25978>.
- (29) Kiyohara, S.; Mizoguchi, T. Searching the Stable Segregation Configuration at the Grain Boundary by a Monte Carlo Tree Search. *J. Chem. Phys.* **2018**, *DETC2018* (1), 241741. <https://doi.org/10.1063/1.5023139@jcp.2018.DETC2018.issue-1>.
- (30) Yang, X.; Zhang, J.; Yoshizoe, K.; Terayama, K.; Tsuda, K. ChemTS: An Efficient Python Library for de Novo Molecular Generation. *Sci. Technol. Adv. Mater.* **2017**, *18* (1), 972–976. <https://doi.org/10.1080/14686996.2017.1401424>.
- (31) Shin, K.; Tran, D. P.; Takemura, K.; Kitao, A.; Terayama, K.; Tsuda, K. Enhancing Biomolecular Sampling with Reinforcement Learning: A Tree Search Molecular Dynamics Simulation Method. *ACS Omega* **2019**, *4* (9), 13853–13862. <https://doi.org/10.1021/acsomega.9b01480>.
- (32) Dong, W.; Wang, H.; He, M.; Ren, F.; Wu, T.; Zheng, Q.; Li, Y. Synthesis of Reactive Comb Polymers and Their Applications as a Highly Efficient Compatibilizer in Immiscible Polymer Blends. *Ind. Eng. Chem. Res.* **2015**, *54* (7), 2081–2089. <https://doi.org/10.1021/ie503645a>.
- (33) Tuckerman, M. E.; Martyna, G. J.; Berne, B. J. Molecular Dynamics Algorithm for Condensed Systems with Multiple Time Scales. *J. Chem. Phys.* **1990**, *93* (2), 1287–1291. <https://doi.org/doi:10.1063/1.459140>.
- (34) Tuckerman, M. E.; Alejandre, J.; López-Rendón, R.; Jochim, A. L.; Martyna, G. J. A Liouville-Operator Derived Measure-Preserving Integrator for Molecular Dynamics Simulations in the Isothermal–Isobaric Ensemble. *J. Phys. Math. Gen.* **2006**, *39* (19), 5629. <https://doi.org/10.1088/0305-4470/39/19/S18>.
- (35) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117* (1), 1–19. <https://doi.org/10.1006/jcph.1995.1039>.
- (36) Kocsis, L.; Szepesvári, C. Bandit Based Monte-Carlo Planning. In *Machine Learning: ECML 2006*; Fürnkranz, J., Scheffer, T., Spiliopoulou, M., Eds.; Lecture Notes in Computer Science; Springer: Berlin, Heidelberg, 2006; pp 282–293. https://doi.org/10.1007/11871842_29.