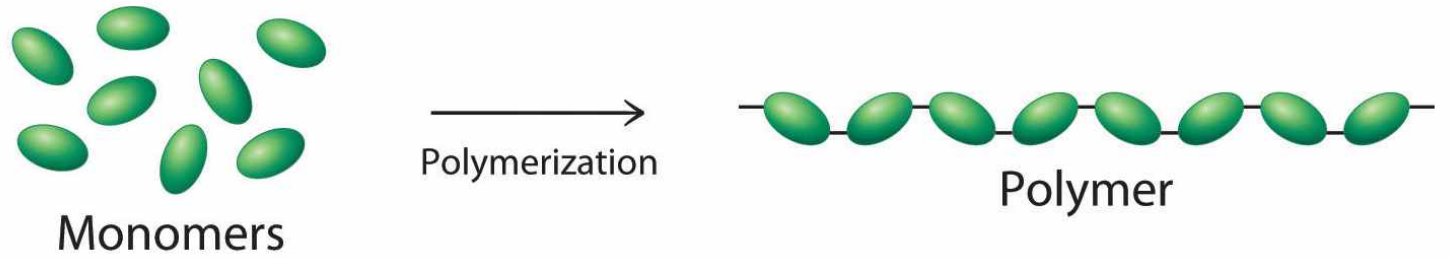
Automated Searches for Inverse Design in Self-Assembly of Branched Block Copolymers

2025. 05. 29

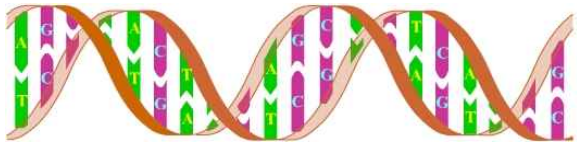
Daeseong Yong
KIAS, Center AI for Natural Sciences

Polymers

- Polymer : a large molecule that is composed of repeating units, called monomers.

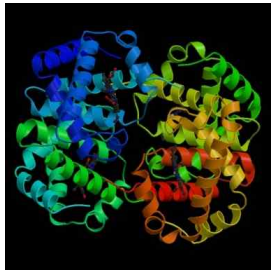


- Examples of polymer



DNA, RNA

Plastics



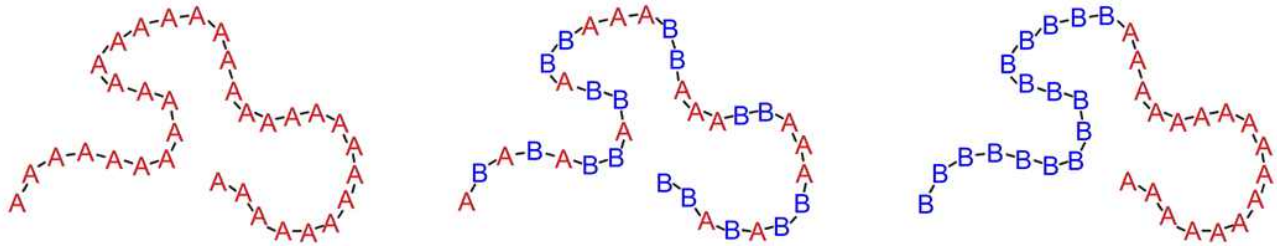
Proteins

Rubbers

https://saylordotorg.github.io/text_general-chemistry-principles-patterns-and-applications-v1.0, M. W. Matsen, 2006

Block Copolymers

- A copolymer is a polymer that consists of two or more distinct monomer units.



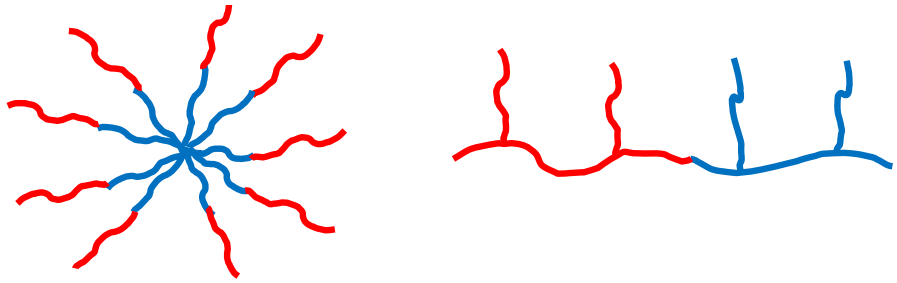
Homopolymer | AB Random Copolymer | AB Diblock Copolymer

- Block copolymers comprise two or more homopolymer subunits linked by covalent bonds.

Linear Polymers

AB Diblock | ABC Triblock

Branched Polymers

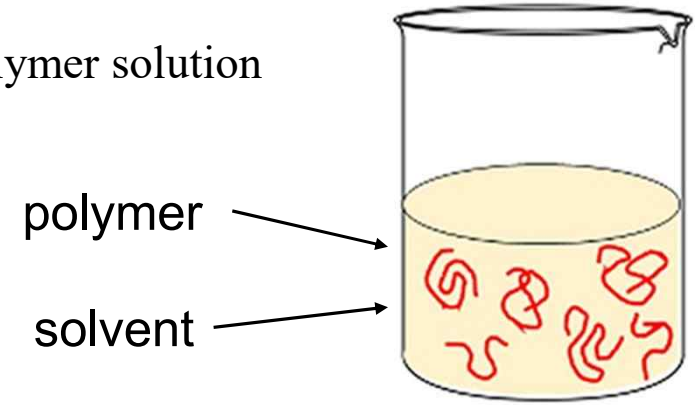


AB multi-arm star-shaped | AB Bottlebrush

Polymeric Liquids

Types of Polymer Liquids

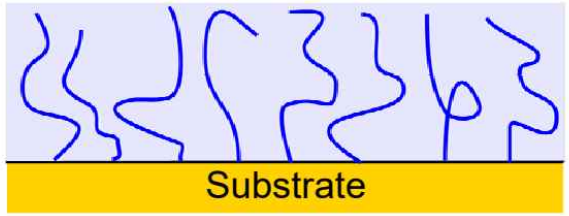
1) Polymer solution



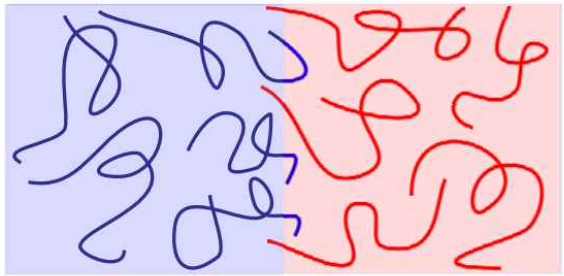
2) Polymer melt (no solvent)

Examples of Polymer Liquids

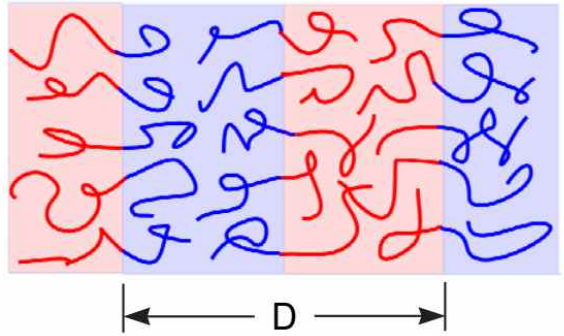
(a) Polymeric Brush



(b) Homopolymer Mixture



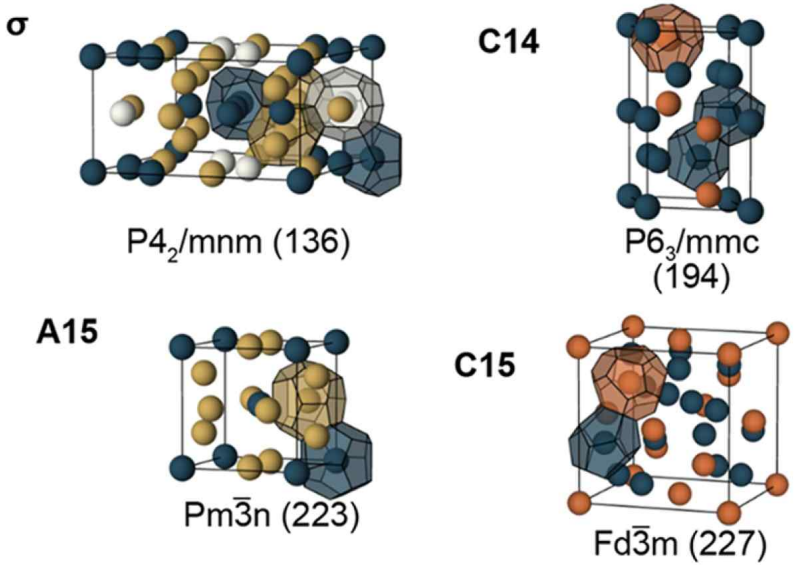
(c) Block Copolymer Melt

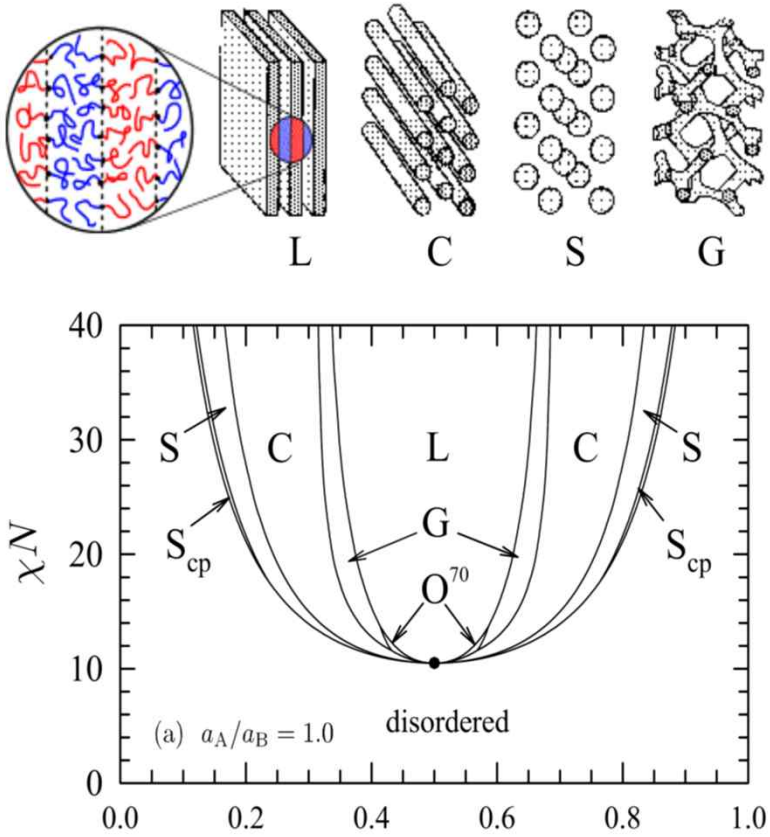


MRS Communications, **11**, 122, (2021),
R. Hoy/University of South Florida; M. Kröger/ETH Zurich
M. W. Matsen, 2006

Self-Assembly of Block Copolymers

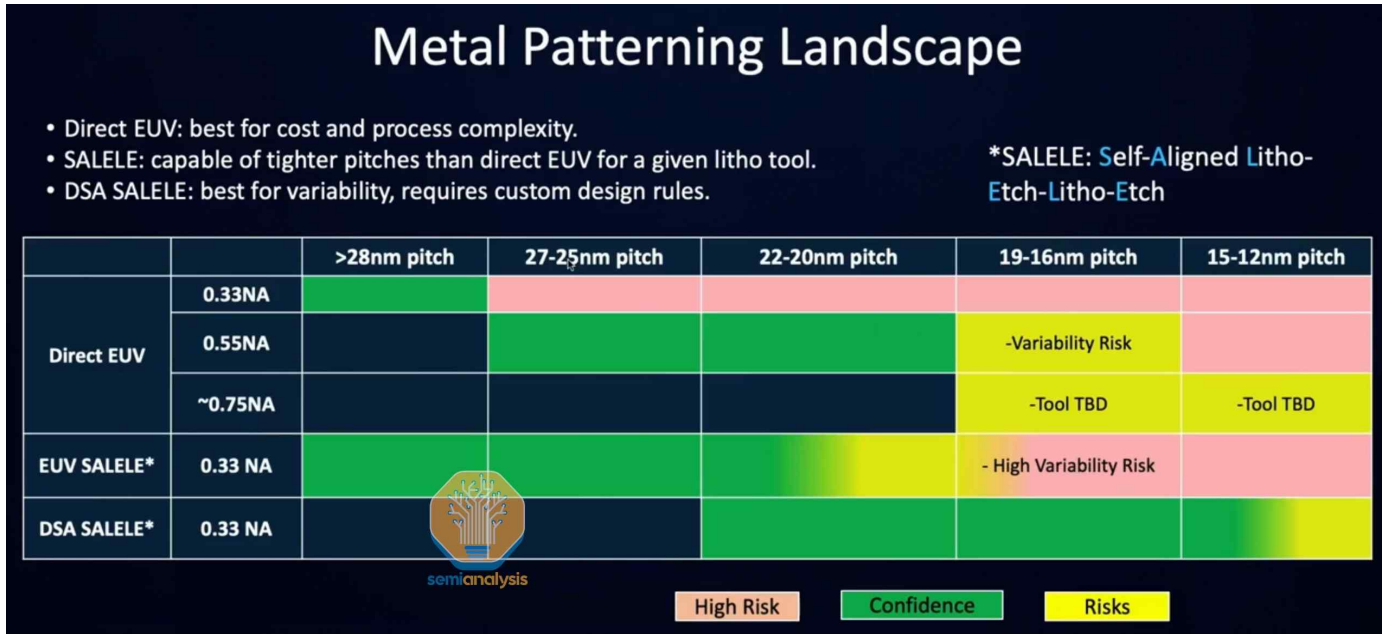
- Block copolymers can spontaneously form various nanometer scale microstructures.
- Depending on interaction strength and polymer architectures, we can produce various stable nano-structures.





M. W. Matsen, *J. Phys.: Condens. Matter*, **14**, R21 (2002)
M. W. Matsen, *J. Chem. Phys.*, **152**, 110901 (2020)
K. D. Dorfman, *Macromolecules*, **54**, 10251 (2021)

Self-Assembly of Block Copolymers



Source: Intel

- They have potential to create novel self-assembled structures with unique properties and diverse applications.

- These complex morphologies offer significant potential for applications in photonics, solar cells, and separation membranes.

https://semianalysis.com/2024/04/18/intels-14a-magic-bullet-directed/#why-directed-self-assembly-is-needed-breaking-the-dose-vs-cd-tradeoff

Statistical Mechanics for Self-assembly

- Statistical mechanics is a tool to investigate the phase behavior of polymers.

- Typical steps for theoretical studies.

1. Consider various nano-structure candidates.

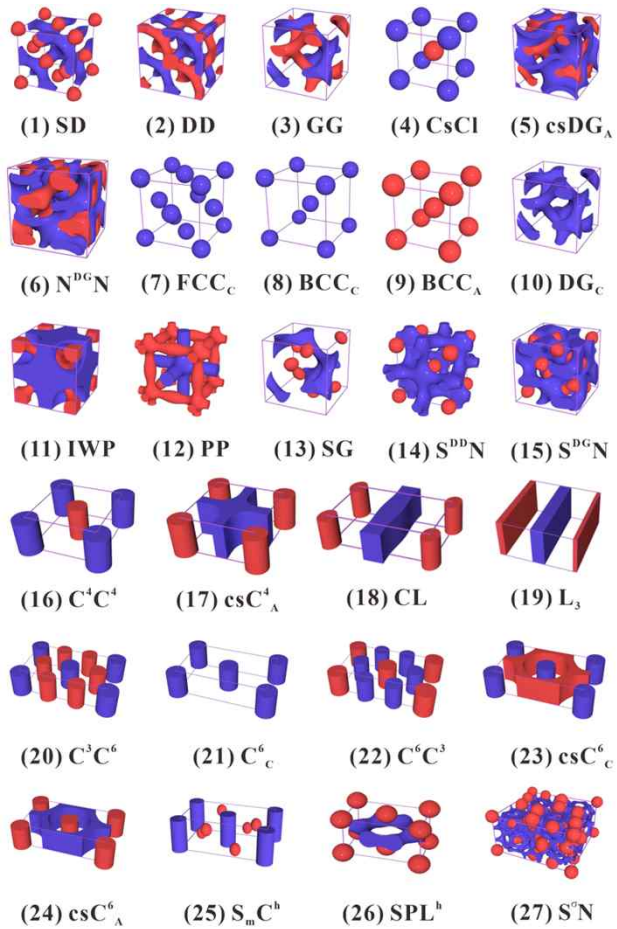
2. Compute the partition function and free energy of each candidate.

$$\mathcal{Z} = \sum_i e^{-\beta E_i}$$

$$\beta F = -\ln \mathcal{Z}$$

i: the index for the microstates that belong to chosen candidate

3. We can identify the stable phase by comparing free energy of the various candidates.

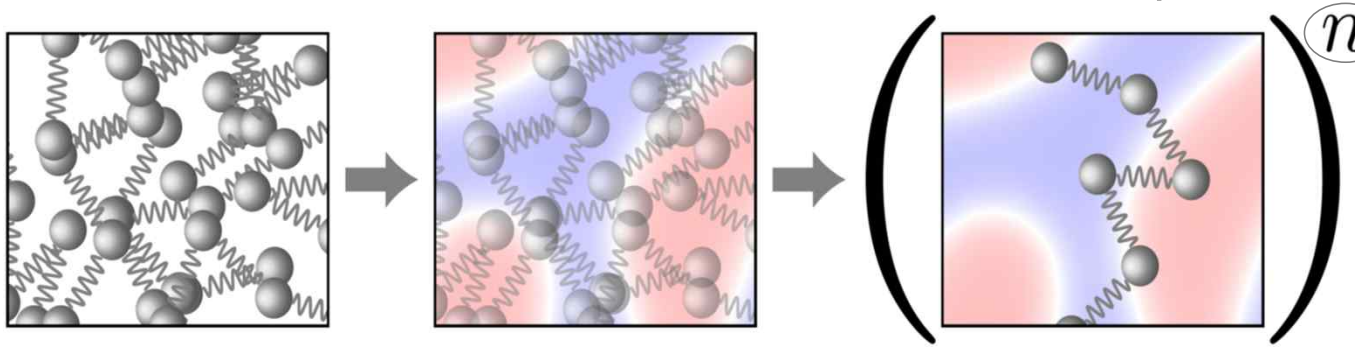


Polymer Field Theory

- In polymer field theory, a many-body problem is converted to the problem of single chain polymer in the potential fields.

Many-Body Problem

Statistical mechanics of single polymer chain in potential fields



Decoupling interactions using the field transformation

Effective Hamiltonian

$$\mathcal{Z} \propto \prod_{j=1}^{n} \int \mathcal{D}\mathbf{r}_j \exp(-\beta U[\mathbf{r}^{nN}]) \quad \Rightarrow \quad \mathcal{Z} \propto \int \mathcal{D}W_+ \int \mathcal{D}W_1 \ldots \mathcal{D}W_{S-1} \exp(-\beta H[W_+, \{W_i\}])$$

Energy of the entire system

AB-type copolymers

$$\mathcal{Z} \propto \int \mathcal{D}W_+ \int \mathcal{D}W_- \exp(-\beta H[W_+, W_-])$$

Self-consistent Field Theory (SCFT)

- There are a few simulation methods based on the polymer field theory, such self-consistent field theory (SCFT) and field-theoretic simulations (FTSs).

- In SCFT, the functional integrals are approximated using the saddle-point approximations,

$$\mathcal{Z} \approx \mathcal{Z}_0 \exp(-\beta H[w_+, \{w_i\}]), \qquad \left.\frac{\delta H}{\delta W_i}\right|_{W_i = w_i} = 0, \qquad \left.\frac{\delta H}{\delta W_+}\right|_{W_+ = w_+} = 0$$

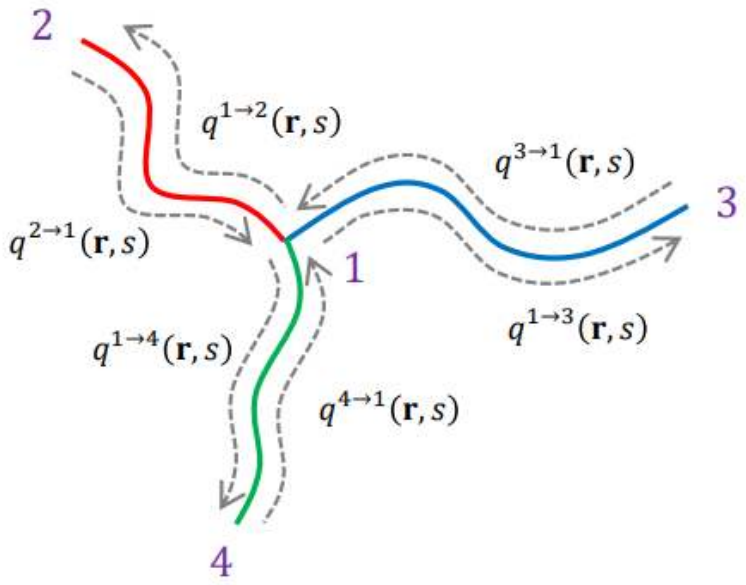
- Self-consistent field theory (SCFT) has been the workhorse method for study of the self-assembly of block polymers.

- Because it is computational efficient (a few seconds ~ a few mins with GPUs), and it allows the direct access of the free energy.

$$F \approx H[w_+, \{w_i\}],$$

Self-consistent Field Theory (SCFT)

- We can obtain the SCFT solution using an iterative method.



Set input fields

$$w_A(\mathbf{r}), w_B(\mathbf{r}),\ w_C(\mathbf{r}), \ldots$$

Compute chain propagators

$$\frac{\partial}{\partial s} q_p^{v\to t}(\mathbf{r},s) = \left[\frac{a^2}{6}\nabla^2 - w_K(\mathbf{r})\right] q_p^{v\to t}(\mathbf{r},s)$$

$$K = K_p^{v\to u}$$

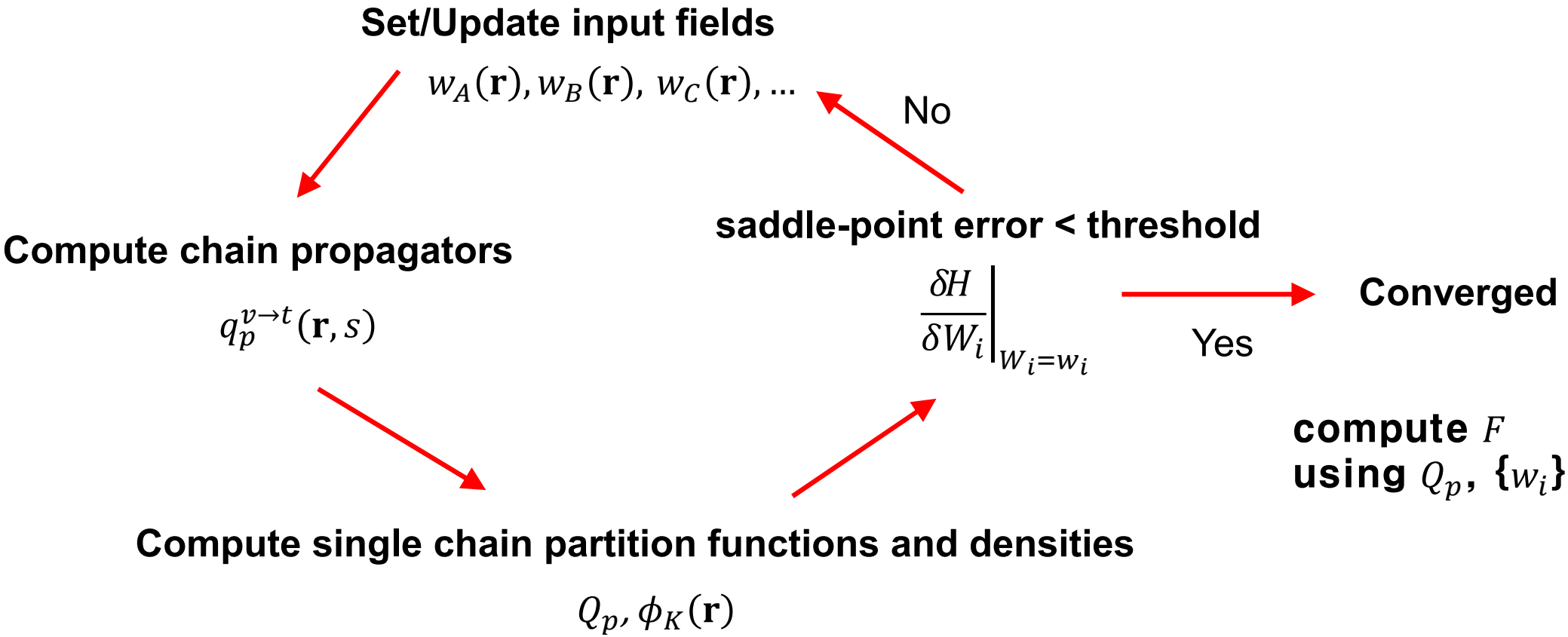
Compute single chain partition functions and densities

$$Q_p = \frac{1}{V}\int d\mathbf{r}\ q_p^{v\to t}(\mathbf{r},s) q_p^{t\to v}\left(\mathbf{r}, N_p^{v\to t} - s\right)$$

$$\phi_K(\mathbf{r}) = \sum \frac{\bar{\phi}_K}{Q_p \alpha_p} \sum_{\substack{(v,t):\\ K_p^{v\to t}=K}} \int ds\ q_p^{v\to u}(\mathbf{r},s) q_p^{u\to v}\left(\mathbf{r}, N_p^{v\to u} - s\right)$$

Self-consistent Field Theory (SCFT)

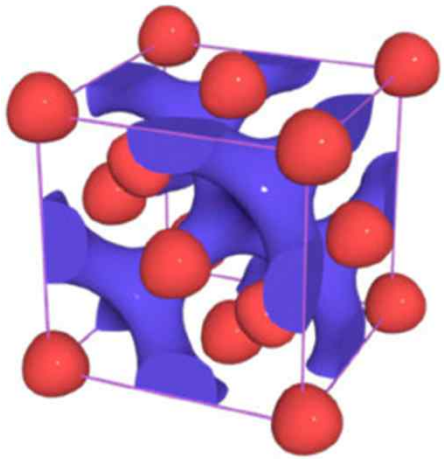
- We can obtain the SCFT solution using an iterative method.



Inverse Design in Self-assembly

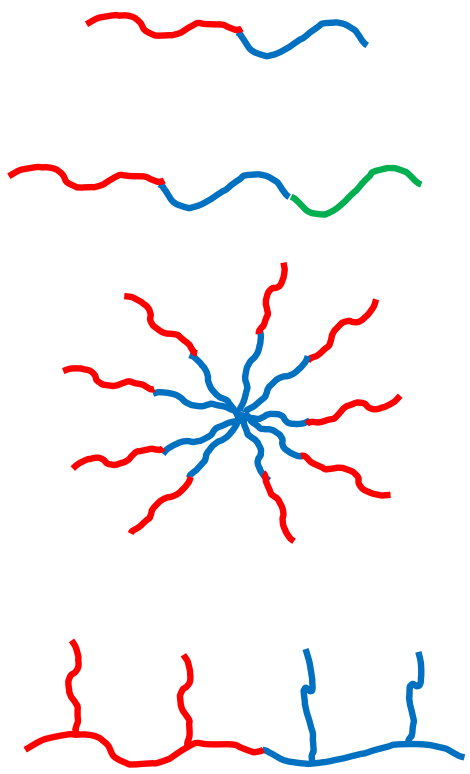
- In the context of self-assembly of block copolymers, the inverse molecular design is finding polymer architectures or systems that stabilize the target phase.

- One of the ultimate goals in block copolymer self-assembly is to develop methods that identify chain architectures capable of stabilizing target nanostructures.



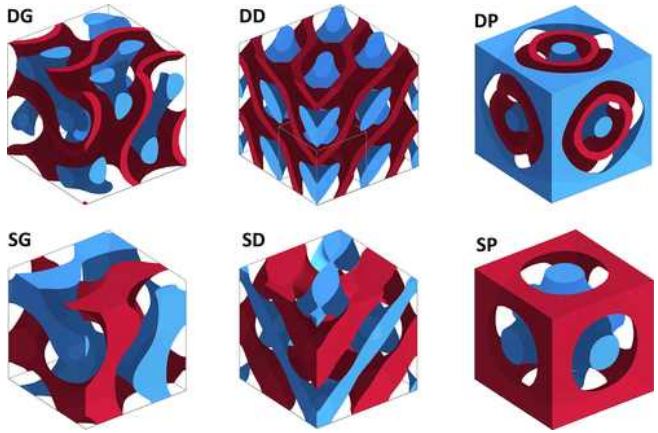
Target Phase

?

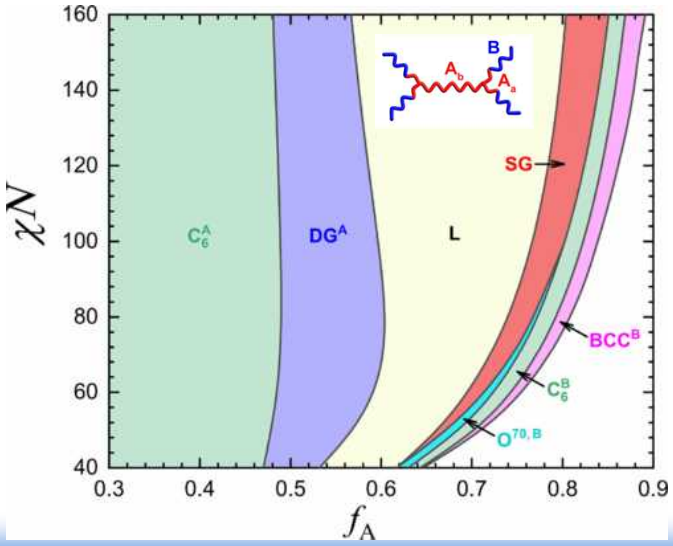


Inverse Design in Self-assembly

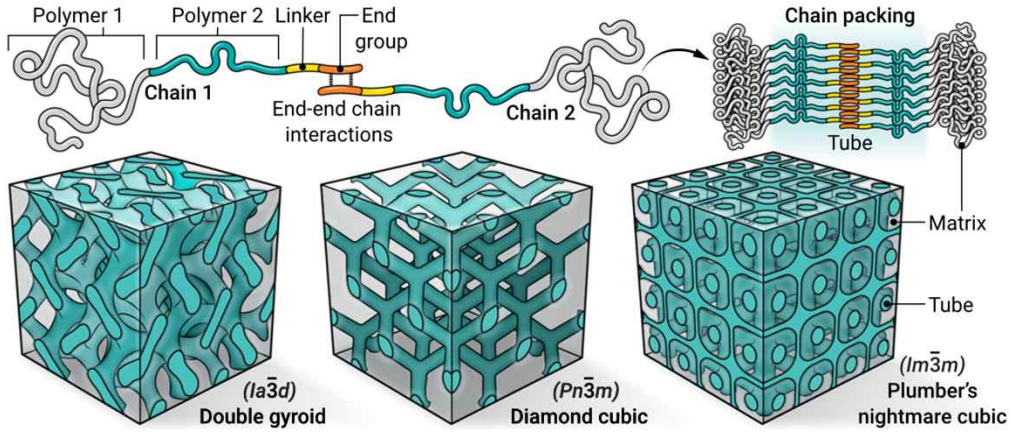
- The network phases, whose domain is continuous, are difficult to synthesize.
- It is worth to publish a paper, if one find good polymer architectures or systems.



Park et al., Phys. Rev. Materials **7**, 105601 (2023)

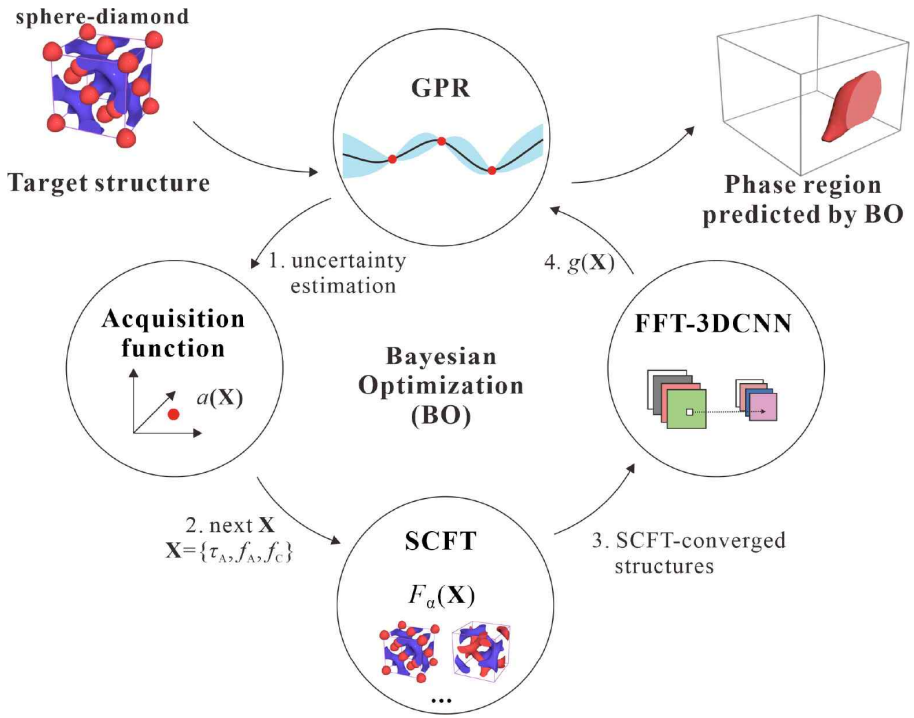


Lee et al., Science 383, 70–76 (2024)



Inverse Design in Self-assembly

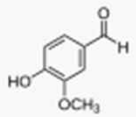
- Recently, Q. Dong et al. employed Bayesian optimization to search for the desired structures self-assembled by ABC-type miktoarm copolymers.



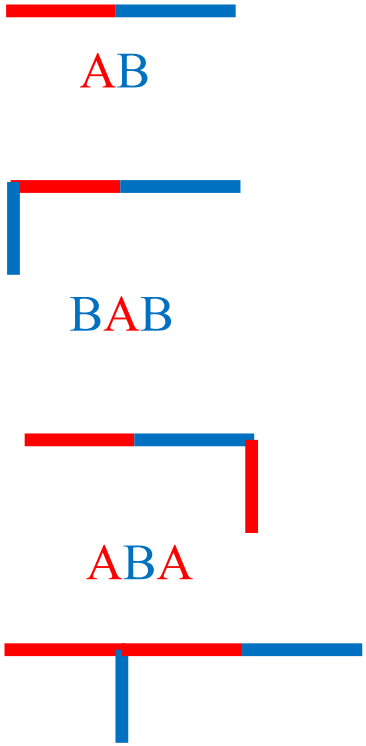
- But, they investigated only one polymer architecture.

Inverse Design in Self-assembly

- Our ultimate goal is to find polymer architectures from arbitrary branched polymers.
- First, we counted the number of polymer architectures using SMILES.
- To uniquely regenerate SMILES, sorting algorithm is added.

Structure	SMILES formula
N≡N	N#N
CH_3–N=C=O	CN=C=O
$Cu^{2+}SO_4^{2-}$	[Cu+2].[O-]S(=O)(=O)[O-]
	O=Cc1ccc(O)c(OC)c1 COc1cc(C=O)ccc1O

Examples)

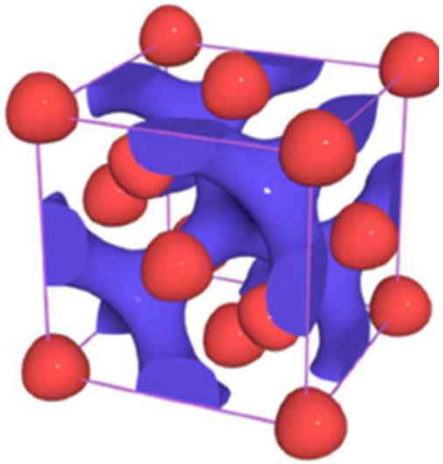


The number of polymer architectures

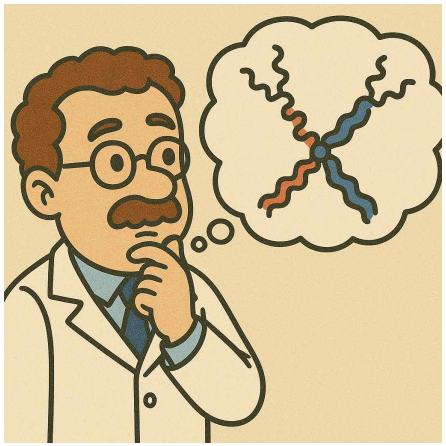
bbcks	# (AB)	# (ABC)
2	1	1
3	4	7
4	10	34
5	36	197
6	106	1061
7	372	6087
8	1249	34966
9	4544	206308
10	16583	

Traditional Research Workflow for Material Discovery

Choose Target

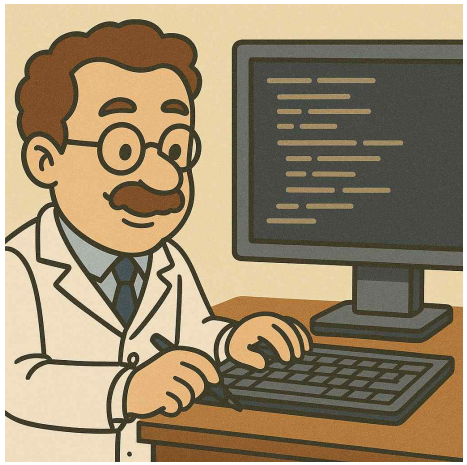


Propose polymers



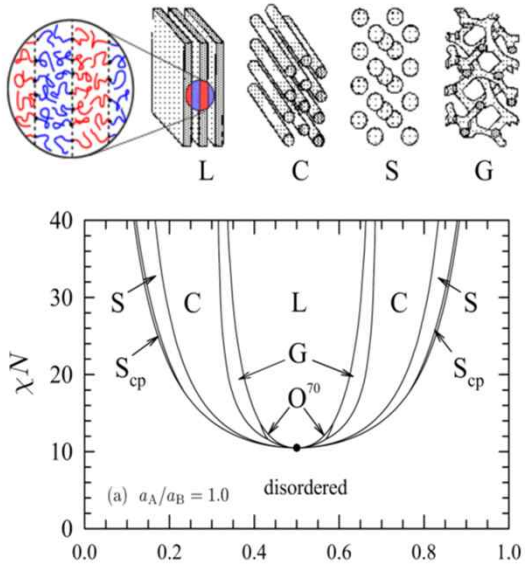
Get hints from experiments and theories

Develop/Update Software



One or two weeks

SCFT computation

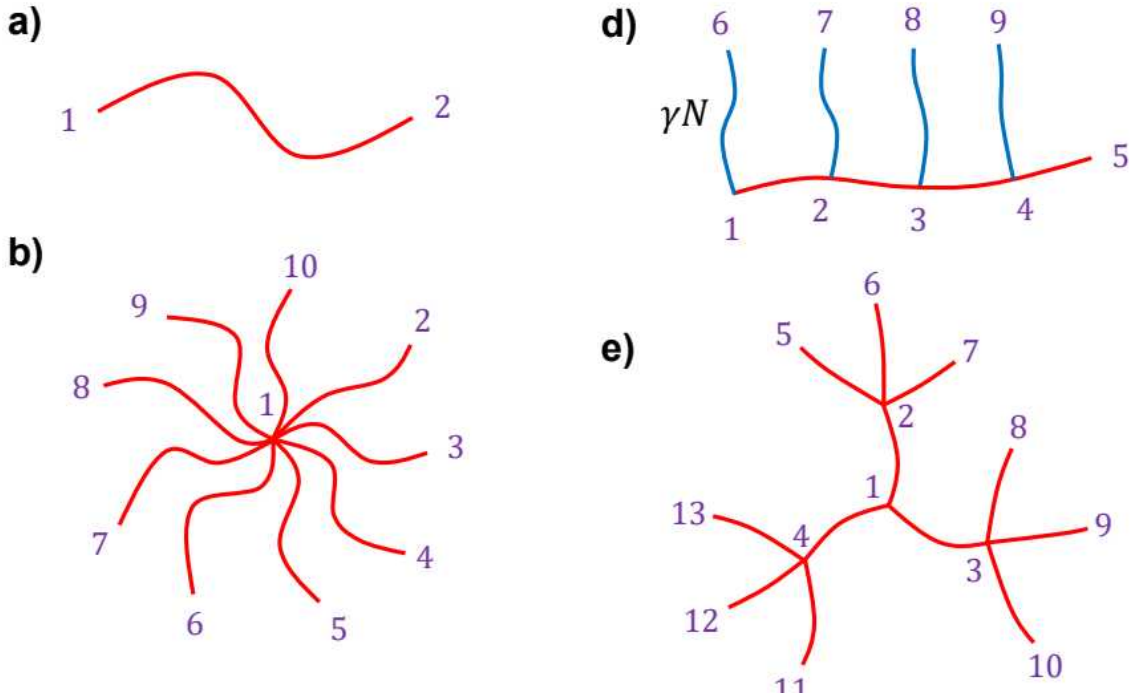


a few hours ~ a few day

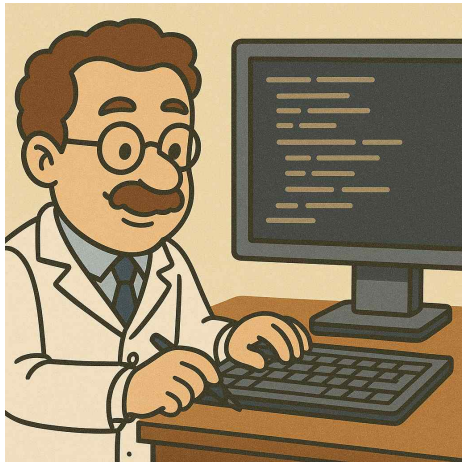
- It wasn't easy to investigate various polymeric systems with this workflow

My Previous Works

- It is hard to develop general software that is efficient for arbitrary branched polymers and mixtures.
- Because each polymer has own strategy to reduce the computational costs, which requires human intervention.



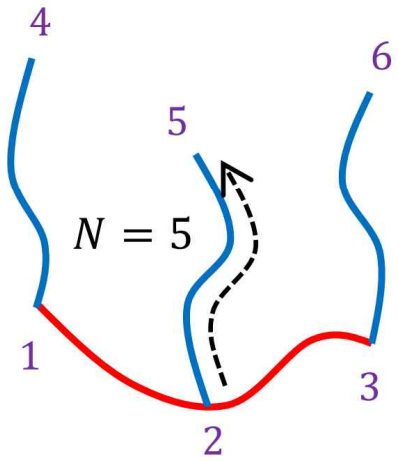
Developing Software



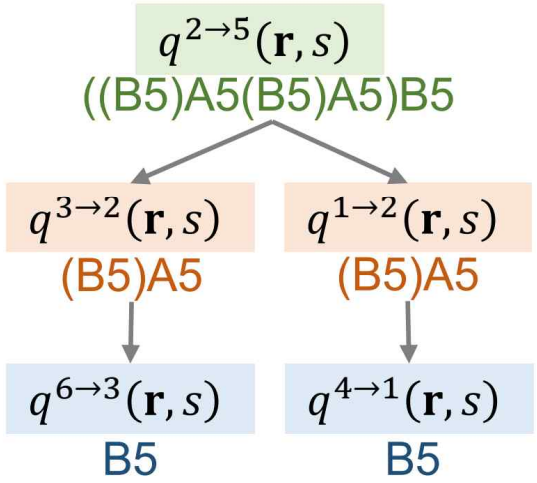
My Previous Works

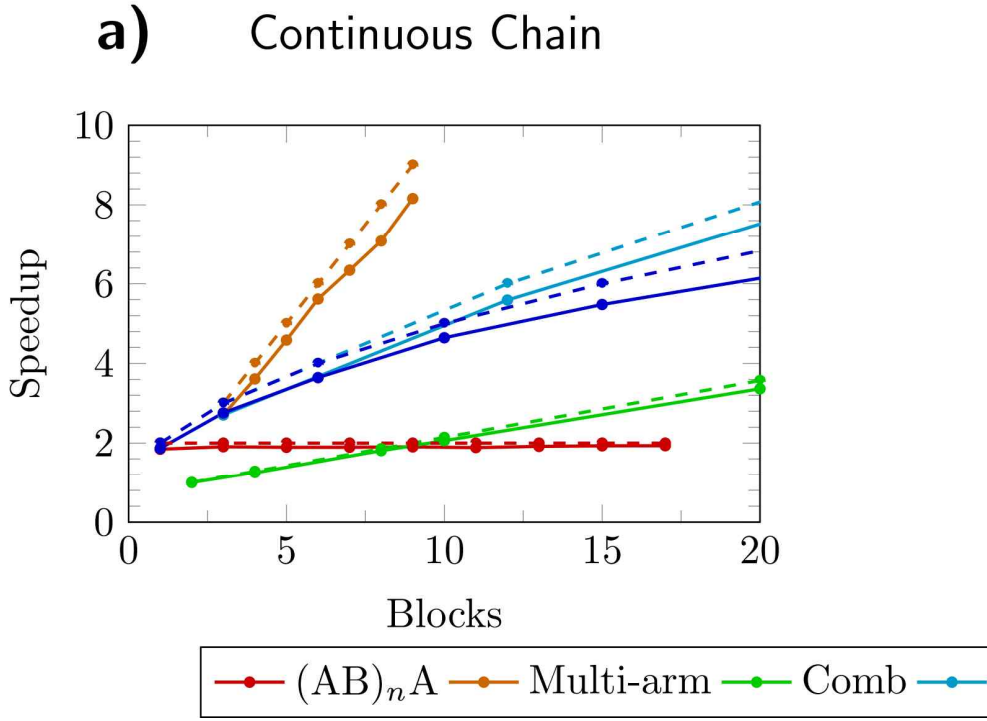
- We developed an algorithmic approach that avoids redundant computations by utilizing the dynamic programming.

Complex Architecture



Dynamic Programming

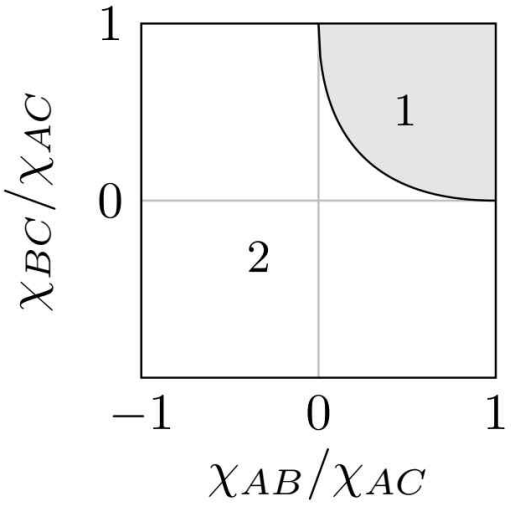




J. Chem. Theory Comput. 2025, 21, 7, 3676−3690

My Previous Works

- Also, we extended our software for multi-monomer systems.



ABC system

$|\chi_{AB}|, |\chi_{BC}| < \chi_{AC}$

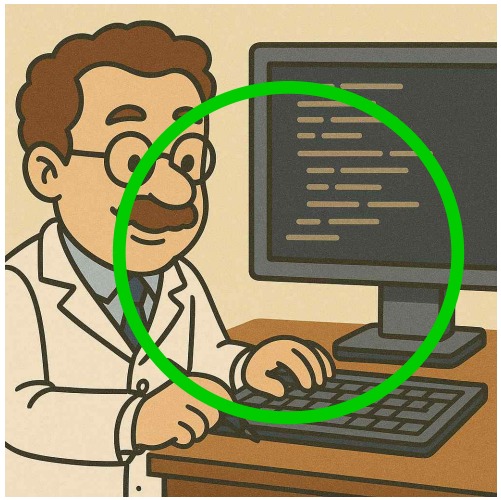
1 or 2 = # of imaginary fields

Polymer Field Theory for Multimonomer Incompressible Models: Symmetric Formulation and ABC Systems

David Morse*, Daeseong Yong, Kexin Chen

Macromolecules 2025, 58, 1, 816–825

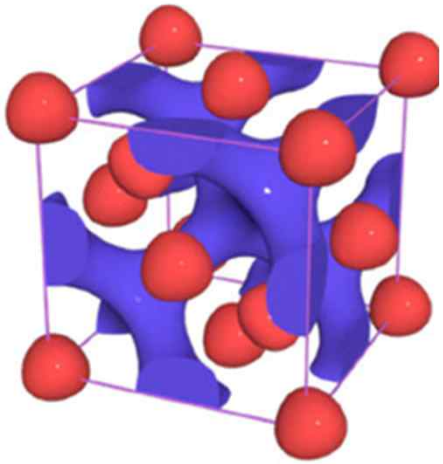
Developing Software



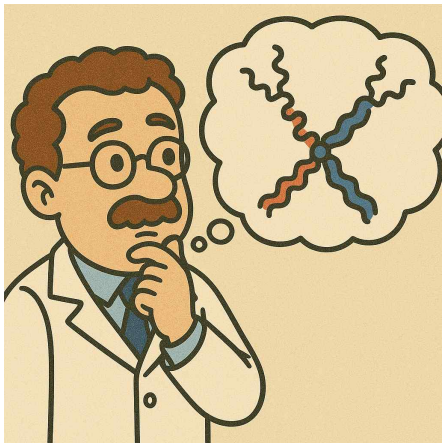
- We developed efficient general software for arbitrary branched polymers and mixtures, thus developing software for each polymeric system is no longer required.

Traditional Research Workflow for Material Discovery

Choose Target

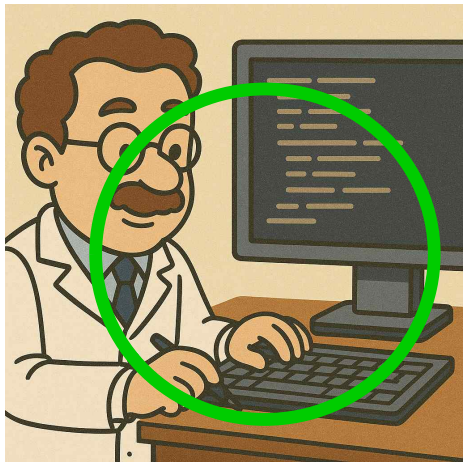


Propose polymers

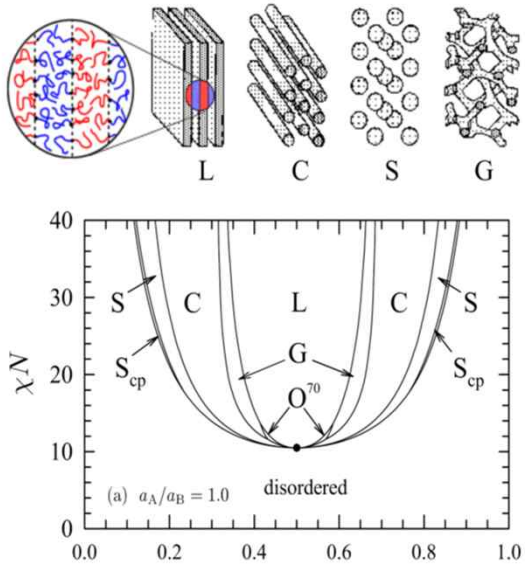


Get hints from experiments and theories

Develop/Update Software



SCFT computation

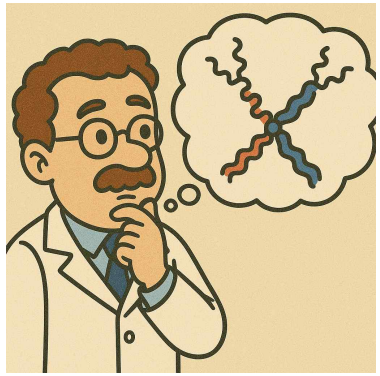


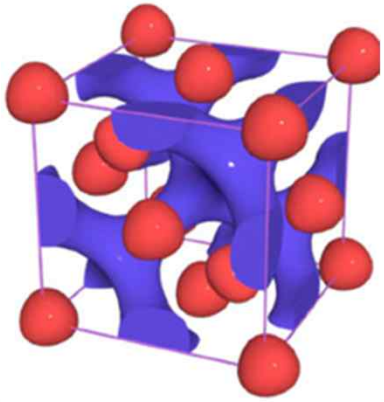
a few hours ~ a few day

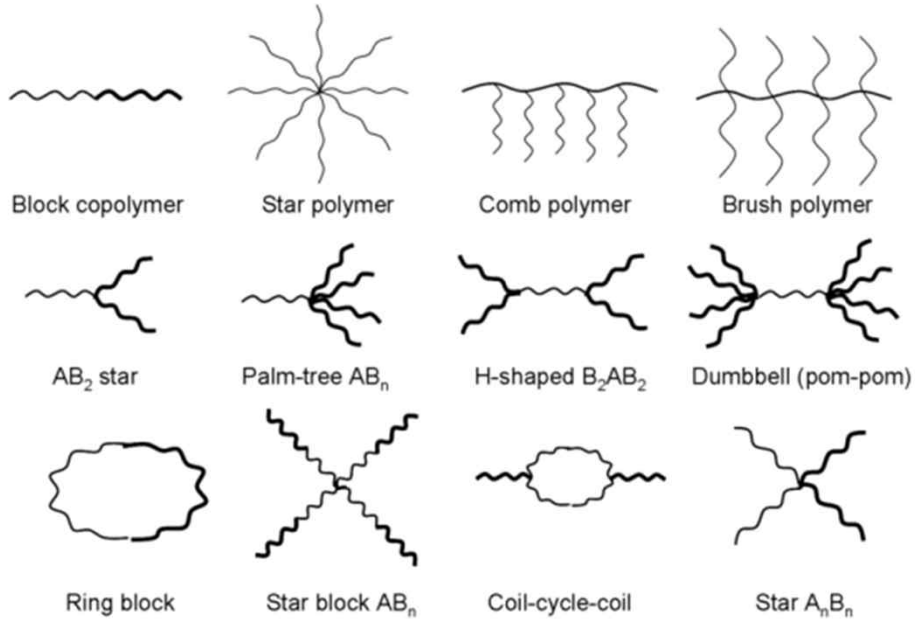
Inverse Design in Self-assembly

- Instead of relying on the physical intuitions and experimental results, we can search all possible candidates.

Propose polymers



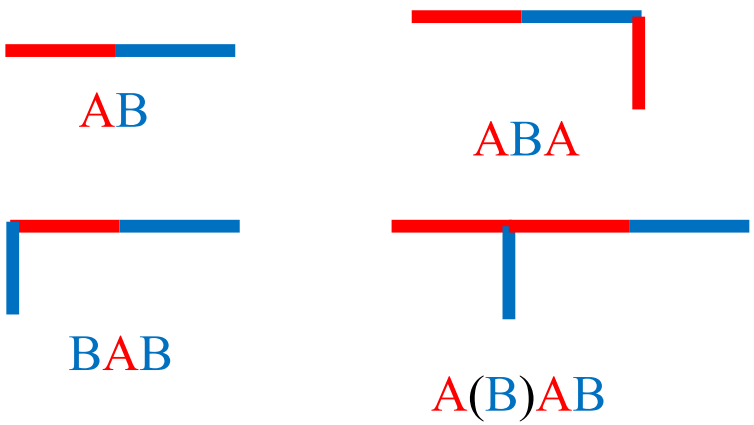




Inverse Design in Self-assembly

- Instead of relying on the physical intuitions and experimental results, we can search all possible candidates.

Examples)



The number of polymer architectures

bbcks	# (AB)	# (ABC)
2	1	1
3	4	7
4	10	34
5	36	197
6	106	1061
7	372	6087
8	1249	34966
9	4544	206308
10	16583	

Inverse Design in Self-assembly

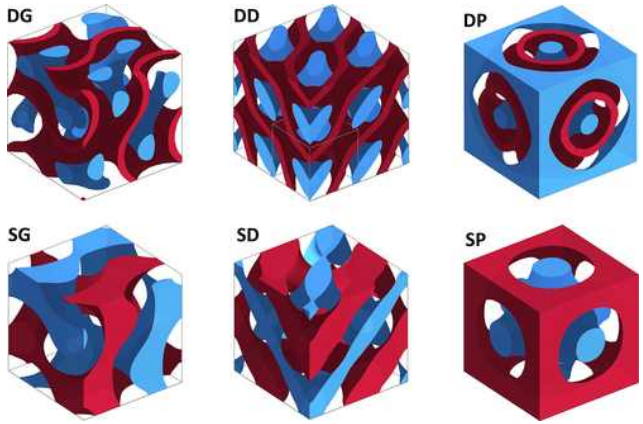
- We are searching for the polymer architectures that stabilize various phases up to $n = 7$.

bbcks	# (AB)
2	1
3	4
4	10
5	36
6	106
7	372

1) Candidate phases.

L, C, BCC, FCC, Simple Cubic, PL, A15, Sigma

Network phases:



2) Parameter Space

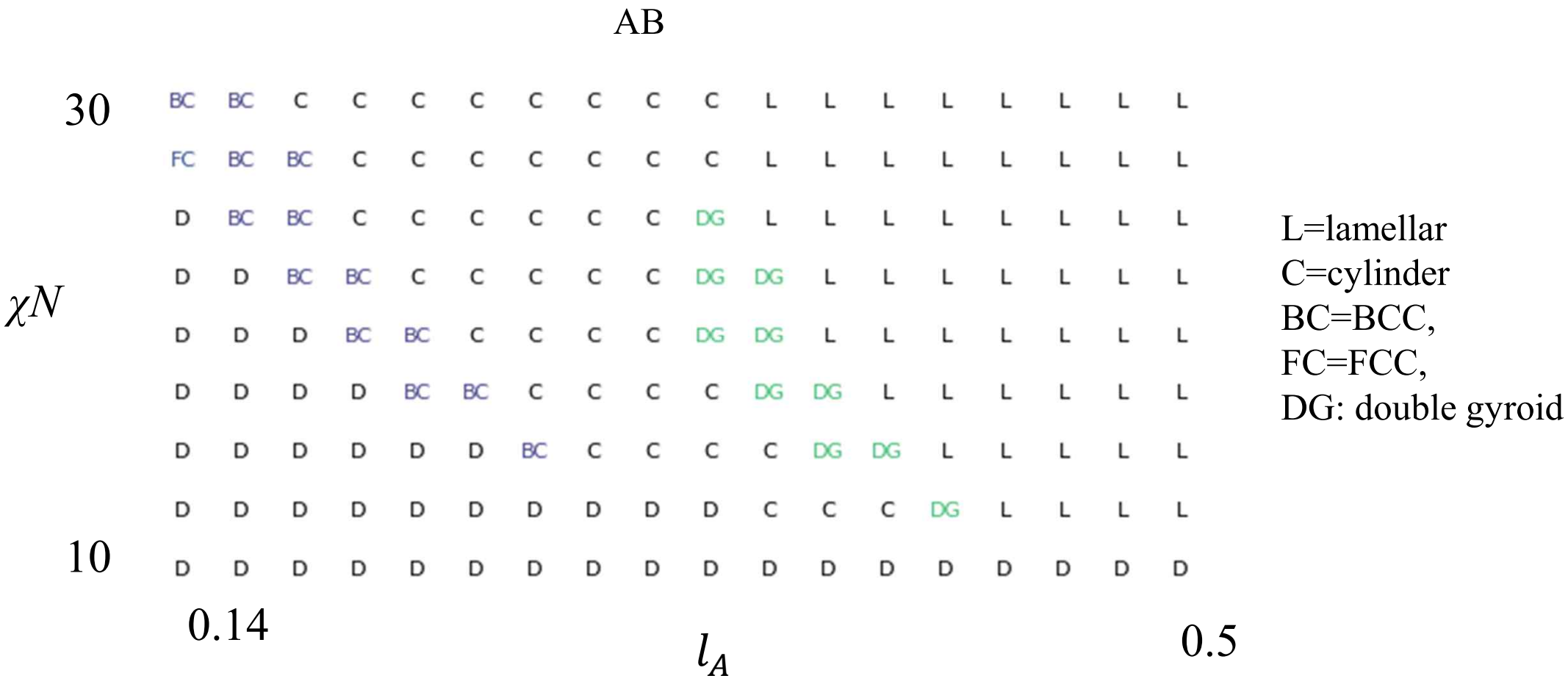
$\chi N \in [10, 30]$ with interval 2.5

$l_A \in [0.14, 0.5]$ with interval 0.02, where l_A is length of A-type block.

$l_B = 1 - l_A$ is length of B-type block.

Inverse Design in Self-assembly

- Example 1)

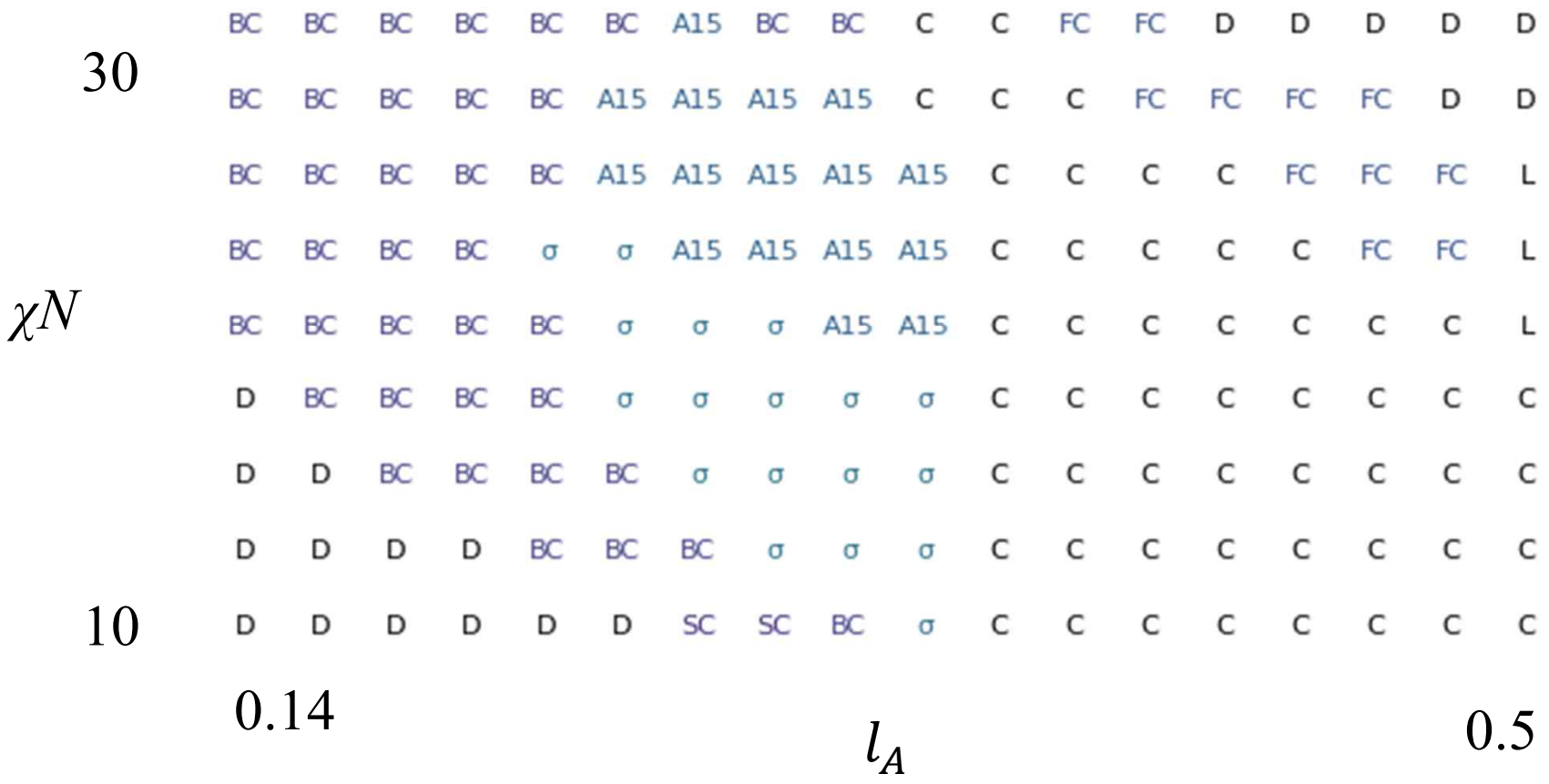


L=lamellar
C=cylinder
BC=BCC,
FC=FCC,
DG: double gyroid

Inverse Design in Self-assembly

- Example 2)

A(A(B)B)A(B)B



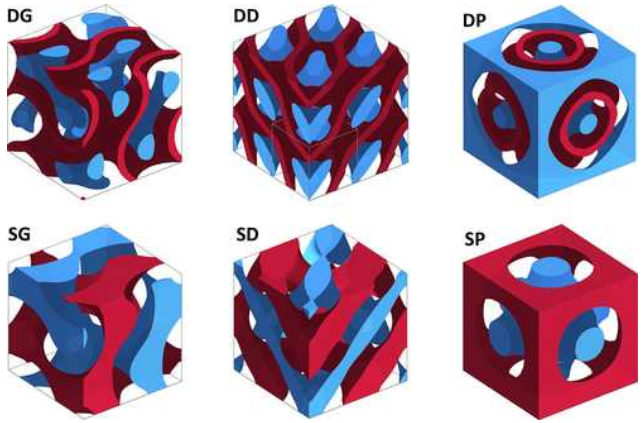
L=lamellar
C=cylinder,
SC=simple cubic
BC=BCC,
FC=FCC,
DG: double gyroid

Inverse Design in Self-assembly

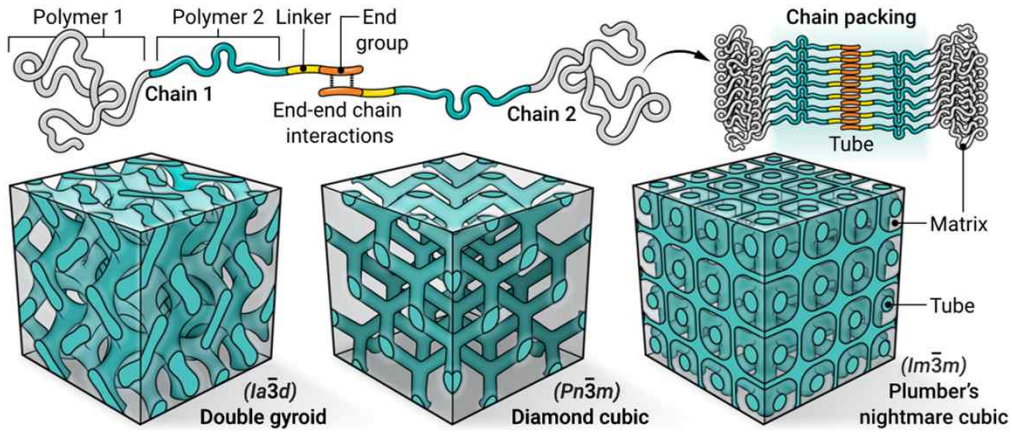
- We found polymer architectures that stabilize each phase.

  Among ~ 300 architectures:

  L, C, BCC, FCC, DG, Simple Cubic, A15, Sigma: **Many**
  DD: **22**
  SG: **17**
  SP: **4**
  SD : **3**
  PL: **2**
  DP (plumbers' nightmare): **0~1**



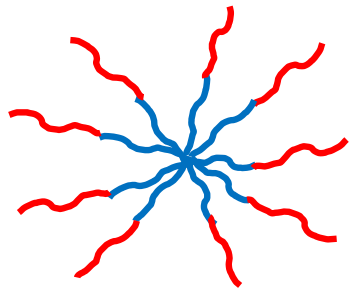
Lee et al., Science 383, 70–76 (2024)



Machine Learning Approach

- Limitations:
  - We investigated only limited number of polymers ~ 300.
  - The parameter windows were small for SD, SP, DP, PL
  - Mixtures are not considered.

- Currently, we are constructing a graph neural network that predicts free energy to extend our approach to more complex architectures.
-
- Data points: ~400k
- Inputs:
  - polymer architecture in graph
  - Interacting parameters, χN
  - Phase
- Output: Free energy



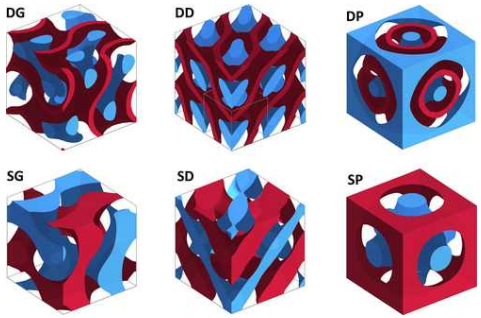
- Its implementation is almost done….

bbcks	# (AB)	# (ABC)
2	1	1
3	4	7
4	10	34
5	36	197
6	106	1061
7	372	6087
8	1249	34966
9	4544	206308
10	16583	

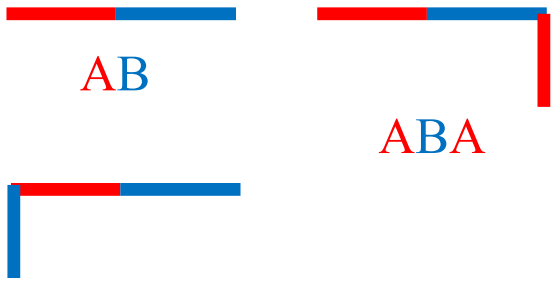
New Research Workflow for Material Discovery

Data Acquisition

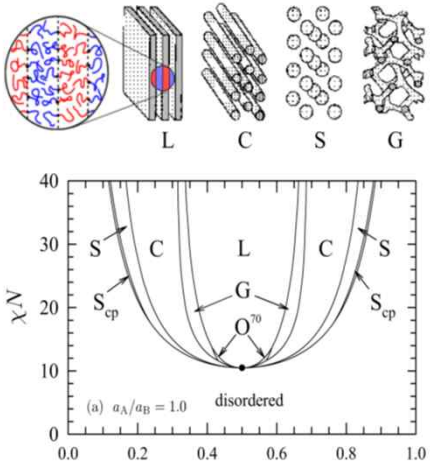
Candidate Phases



Various Architectures, Interaction parameters



SCFT computation



After Training

Candidate Phases

All possible Architectures, Interaction parameters

SCFT computation by ML

Summary

- We have developed efficient software that automatically eliminates redundant computations in self-consistent field theory (SCFT) computation.
- Using this approach, we investigated the self-assembly behavior of AB-type branched block copolymers in brute force manner.
- This approach is currently being extended to more complex branched polymers using graph convolutional networks

Thank you