

2023年度 第2回計算科学フォーラム

# Developments and Applications in Multiscale Molecular Dynamics Simulations for Biomolecular Condensation Using GENESIS

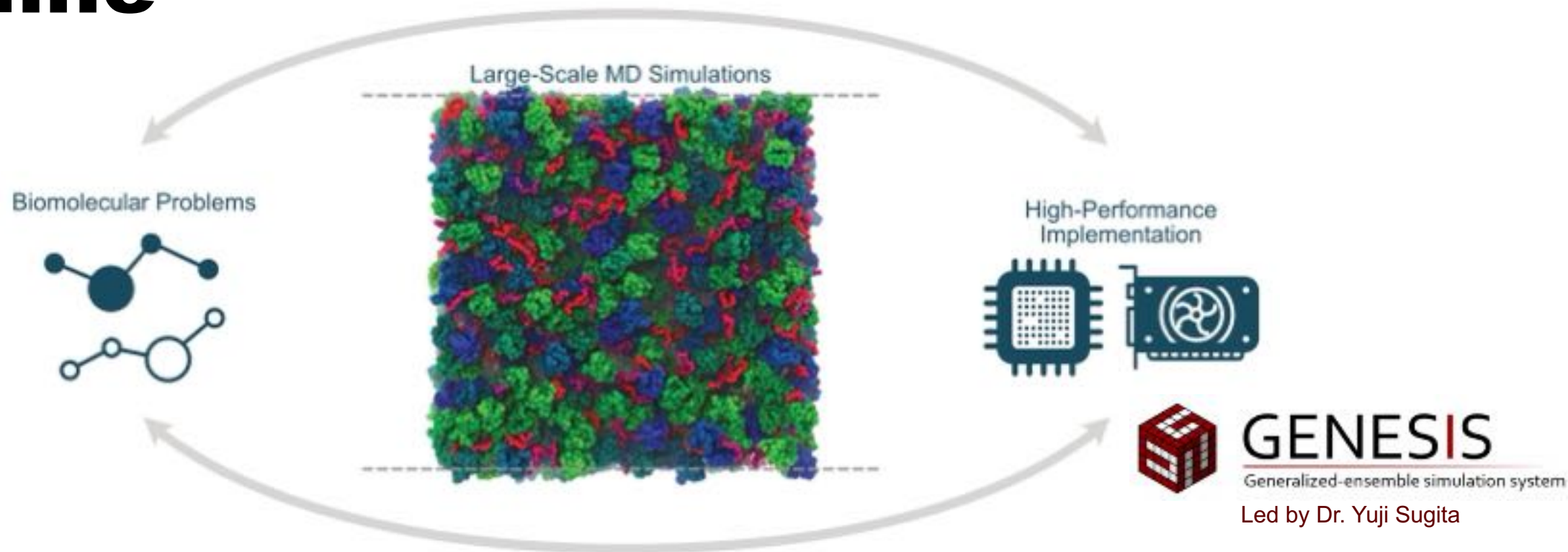
---

Cheng Tan

Computational Biophysics Research Team, RIKEN Center for Computational Science

2024-03-22 Friday

# Outline



- **Developments: A User-Friendly, High-Performance CG MD Software.**

*C.Tan et al. JCP, 2020; PLoS Comput. Biol., 2022. J. Jung, C.Tan, Y. Sugita, in revision, 2024.*

- **Applications: A Comprehensive Exploration of the Regulation of LLPS.**

*C.Tan, A. Niitsu, Y. Sugita, JACS Au, 2023.*

# Liquid-Liquid Phase Separation (LLPS) in Biology

Physics



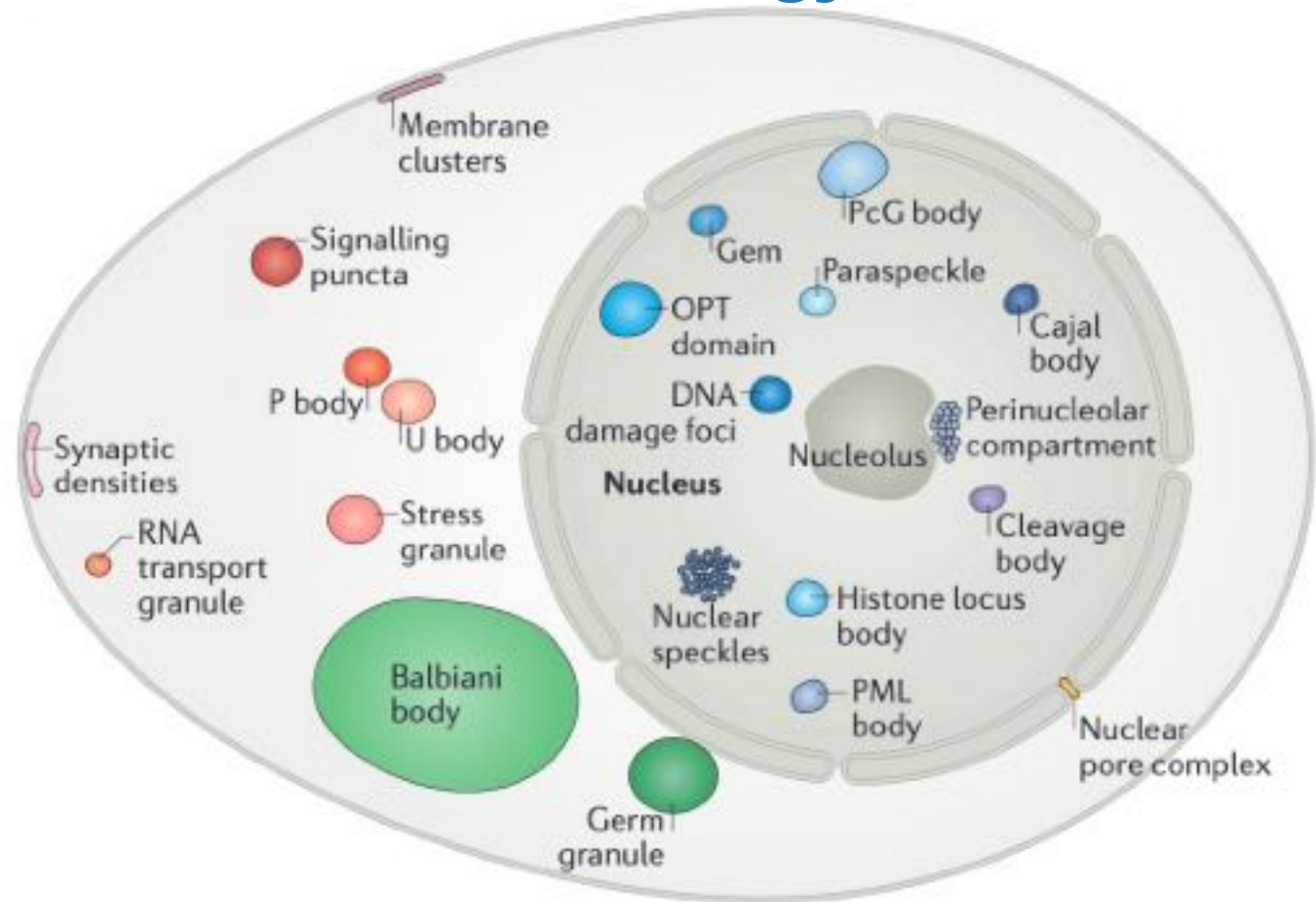
Biology

What is LLPS? 「相分離」



Olive oil with Balsamic Vinegar

Wikipedia, by Leon Brocard



Brangwynne CP, Hyman AA, et al. 2009 *Science*

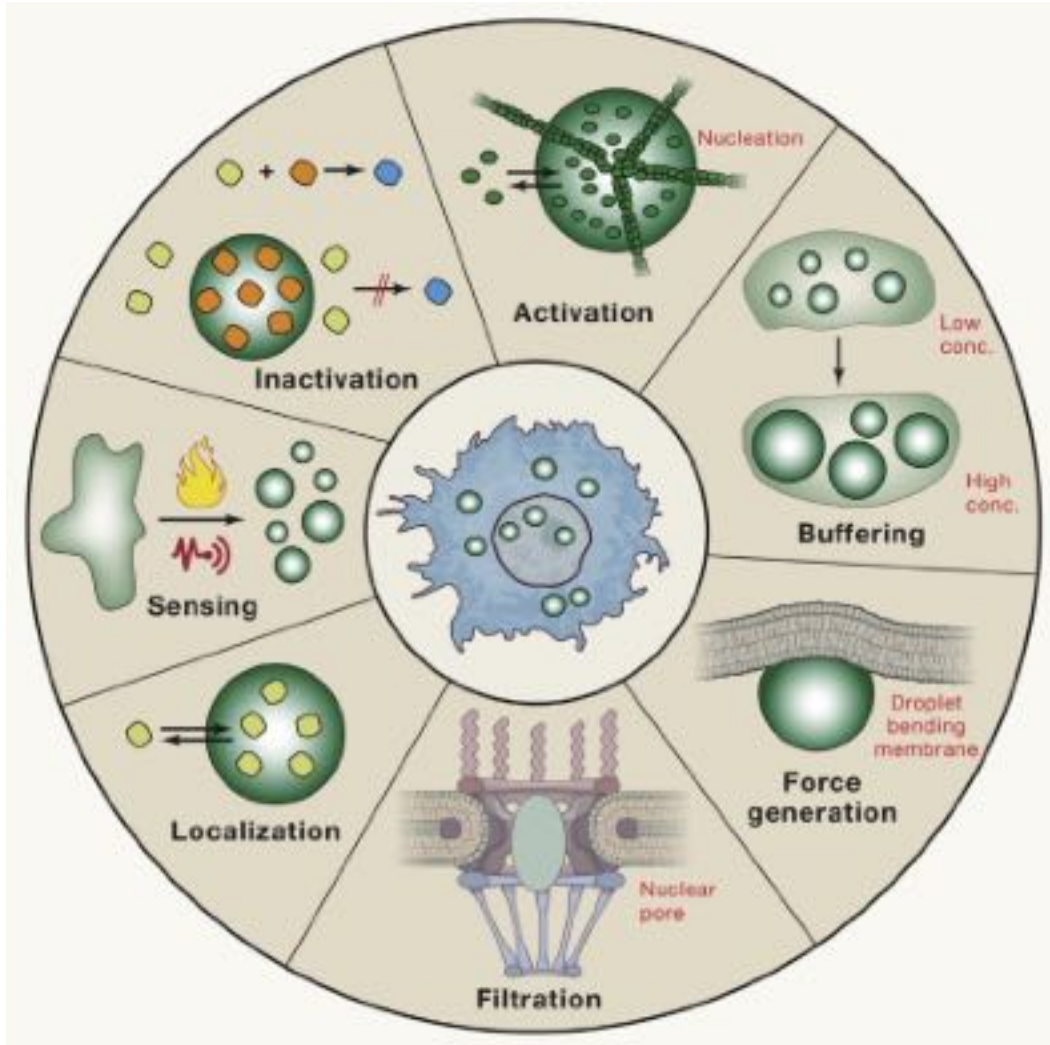
Banani et al. 2017 *Nature Rev. Mol. Cell Biol.*

Same thing, different names:

“droplet”, “condensate”, “granule”, “membraneless organelle”

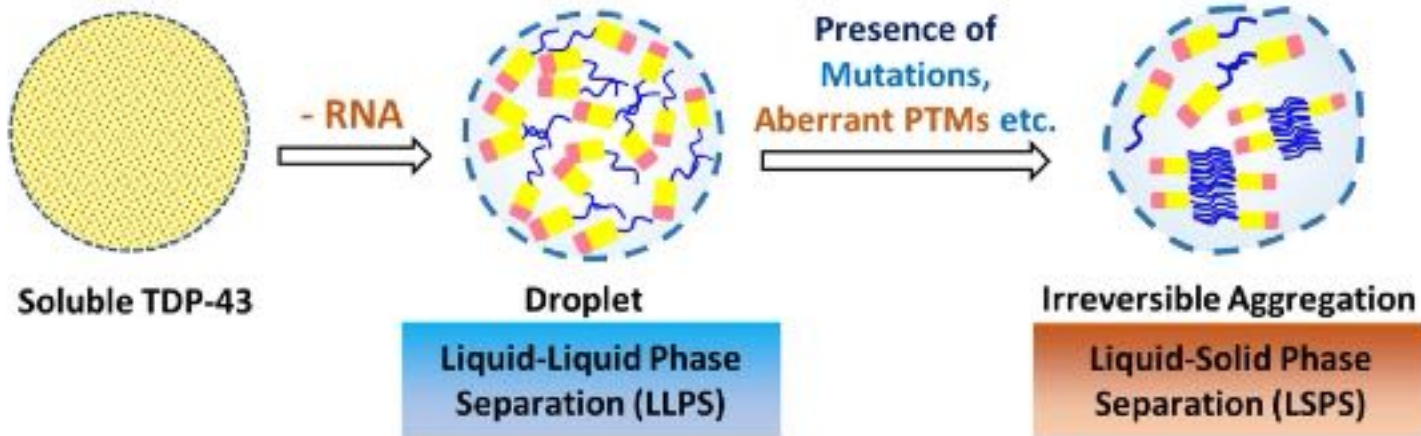
# Biological Problem: Condensation Regulation

## “Good” phase separation

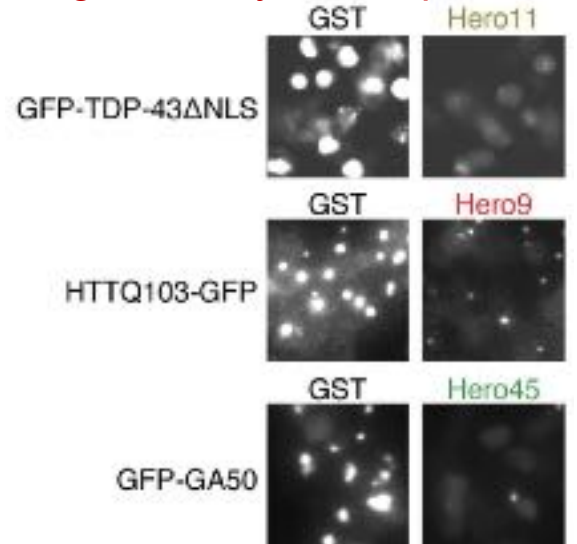


S. Alberti et al. 2019, Cell.

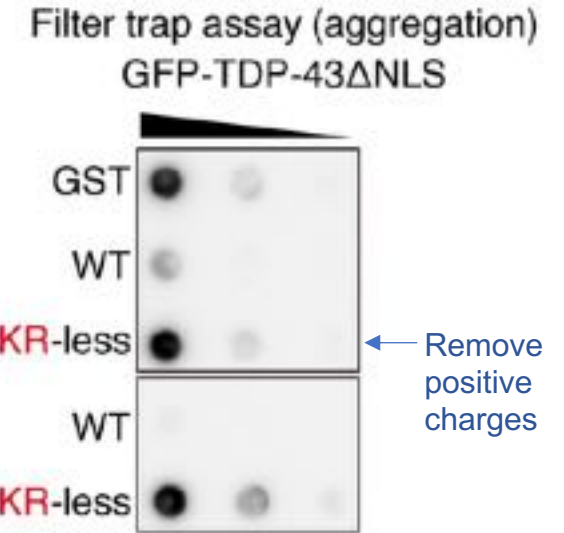
## “Bad” phase separation



## Regulation by “Hero” proteins



Tsuboyama et al. 2020 PLoS Biol.

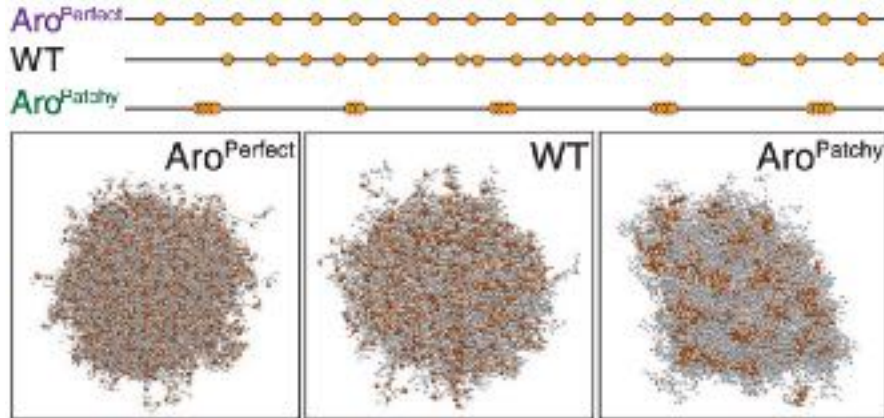
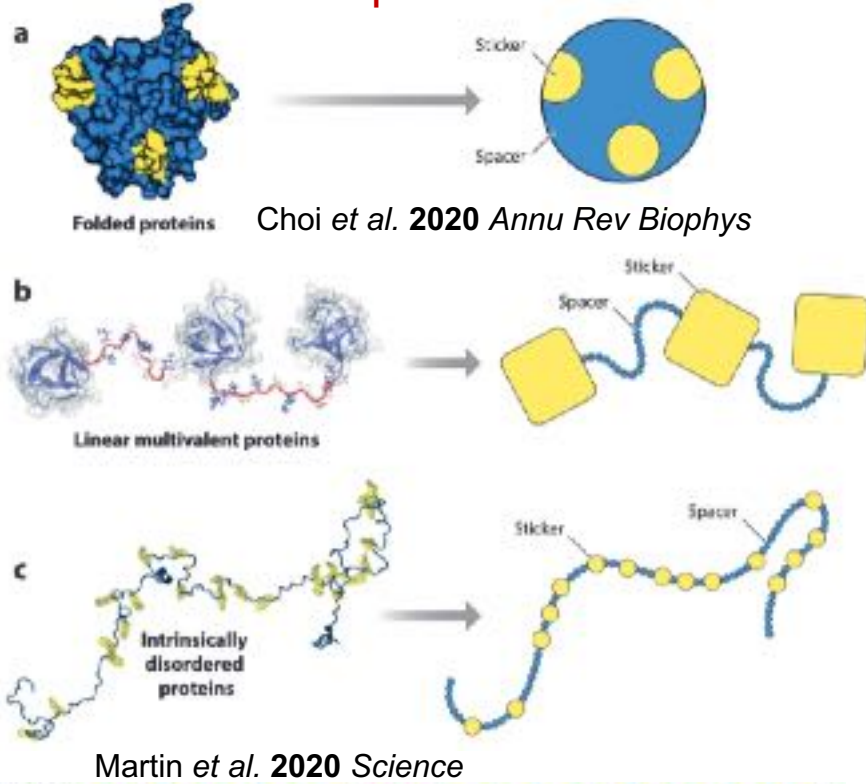


**Question 1:** What's the driving force of the passive regulation?

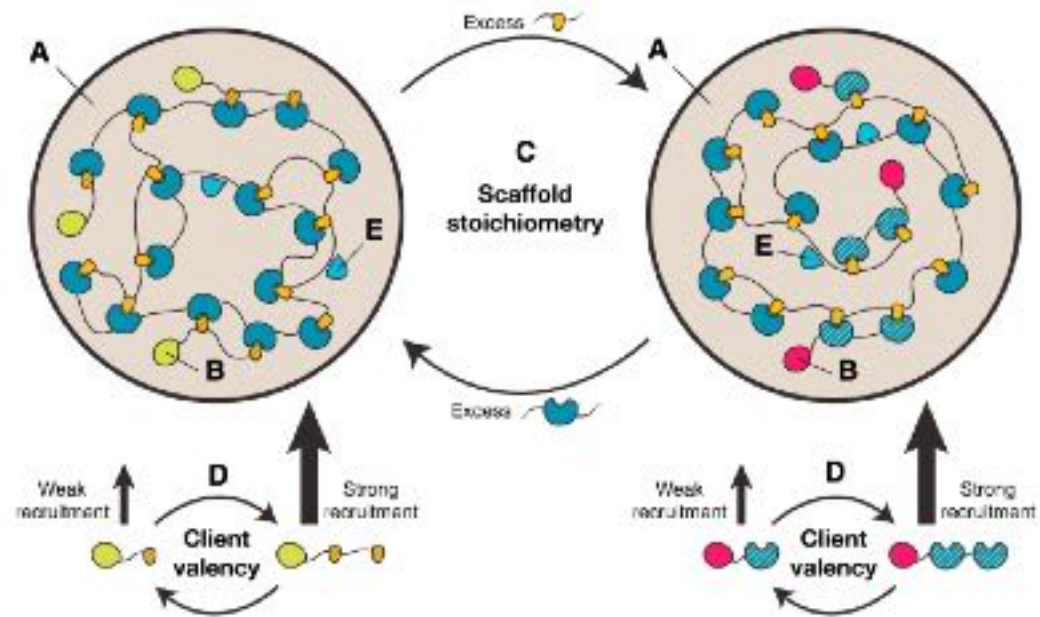
0: Background | 1: GENESIS CG Models | 2: MD of Biomolecule LLPS

# Theoretical Frameworks of Biomolecular LLPS

## “Stickers-and-spacers”



## “Scaffold-and-client”



Clients are not always facilitators

Most existing theories only talk about attractive interactions!

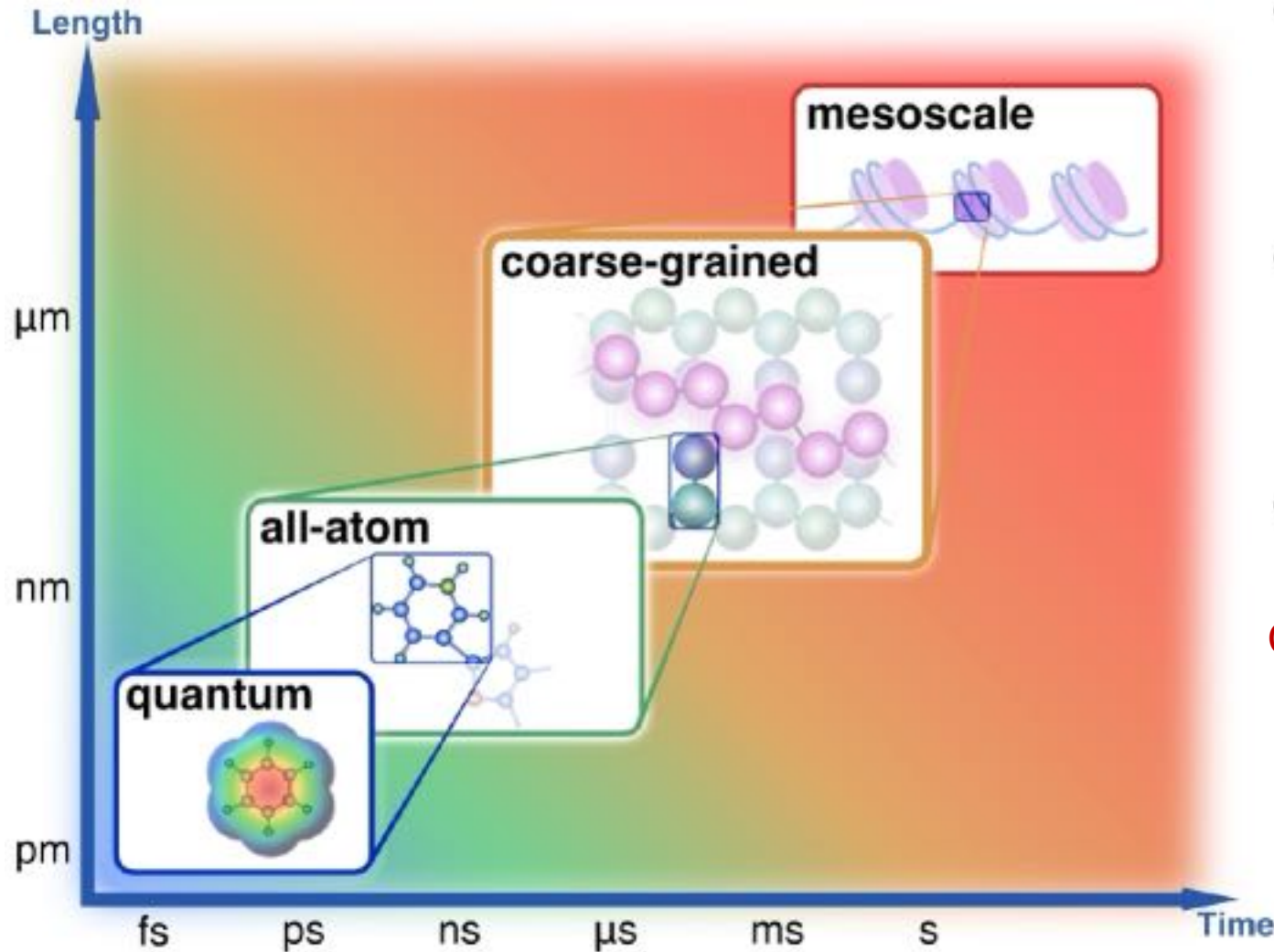
**Q**uestion 2: What's the role of repulsive interactions in the regulation of LLPS?

# Multi-scale Simulations of Biomolecular Systems

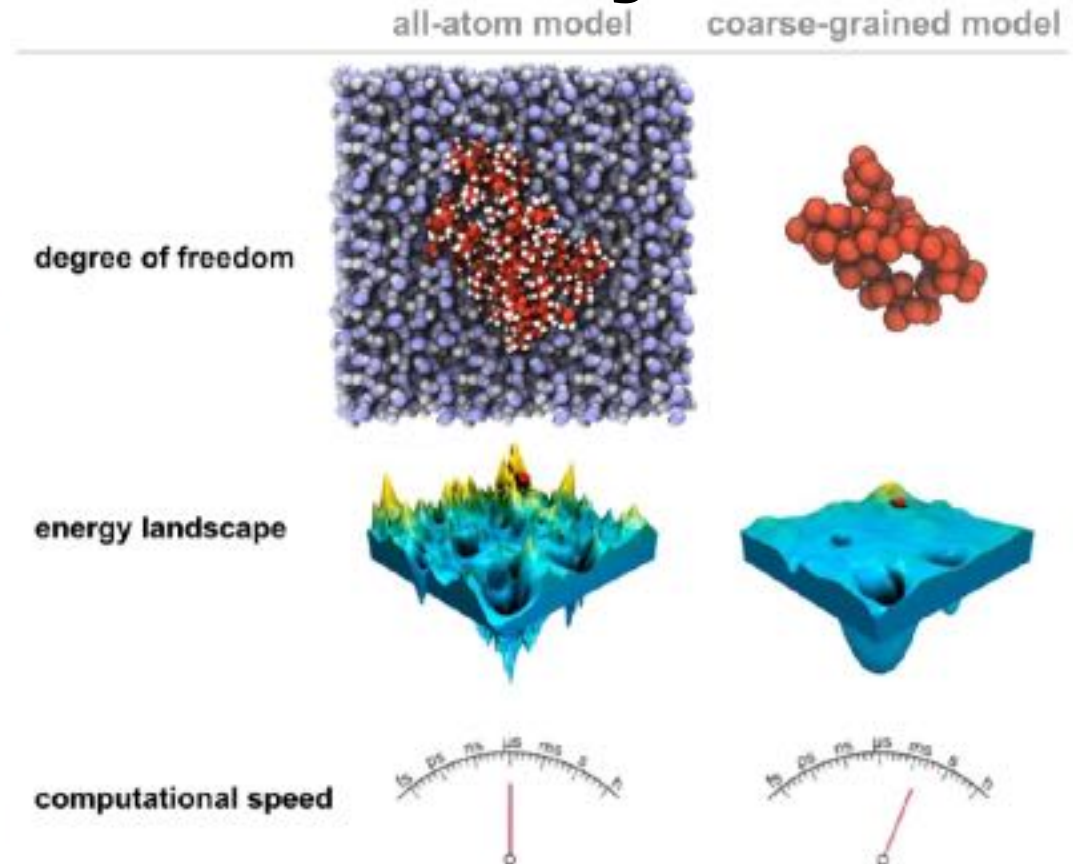
Task of MD: numerically solving Newton's equations of motion

$$M\ddot{X}(t) = -\nabla U(X) + f(X, \dot{X})$$

Potential energy  $U(X)$  determines accuracy and efficiency.

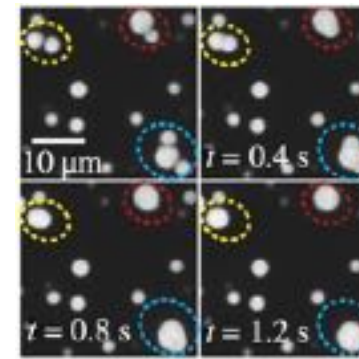


Adapted from A. Kolinski *et al.* 2016, *Chem. Rev.*



A. Kolinski *et al.* 2016, *Chem. Rev.*; S Takada *et al.* 2015 *Accounts Chem. Res.*

Choose a suitable model for target system



## Liquid-Liquid Phase Separation (LLPS)

- Length-scale: 0.1-10 μm
- Time-scale: ms-s
- Number of proteins:  $10^3$ - $10^5$

C.D. Crowe *et al.* 2018, *Interface Focus*

# Popular Coarse-grained Models for Biomolecules

Residue-level coarse-graining: ~10 atoms / CG particle

- Protein: AICG2+

W. Li *et al.* 2014, *PNAS*.

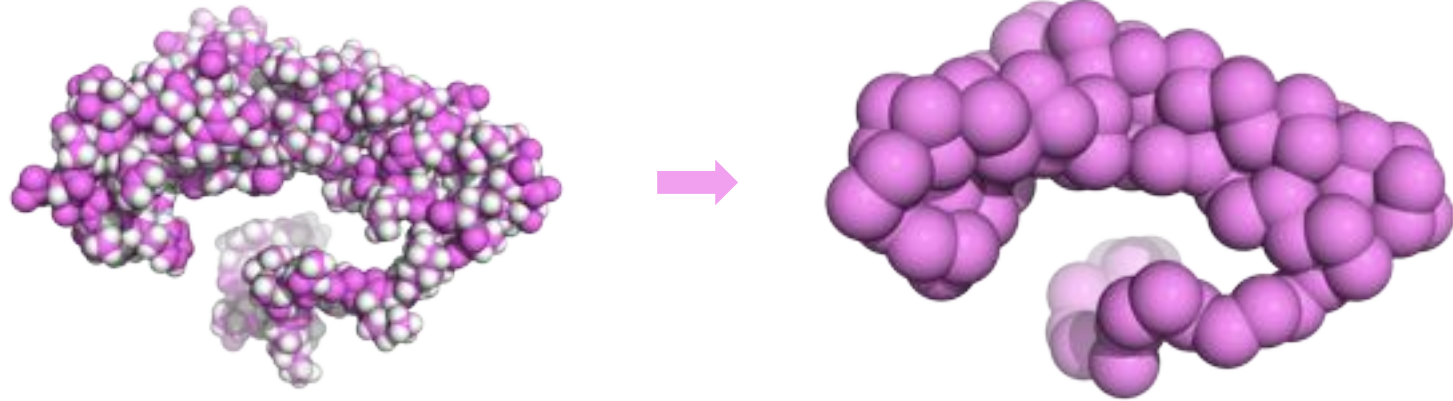
- DNA: 3SPN.2C

G. Freeman *et al.* 2014, *JCP*.

- RNA: Go-like

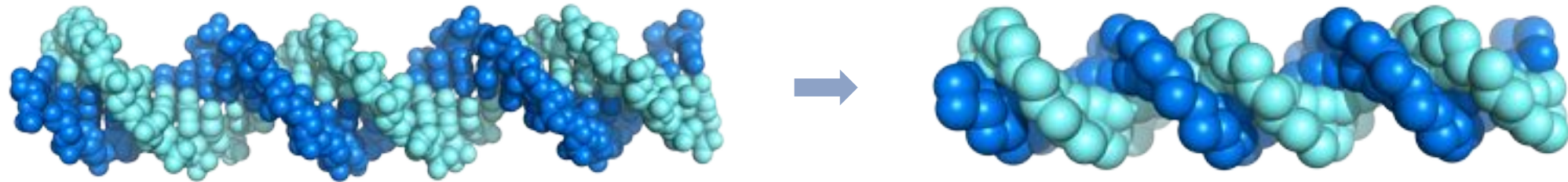
N. Hori *et al.* 2012, *JCTC*.

**1 bead / amino acid**



**3 beads / nucleotide:** Phosphate, Sugar, Base

**3 beads / nucleic acid**

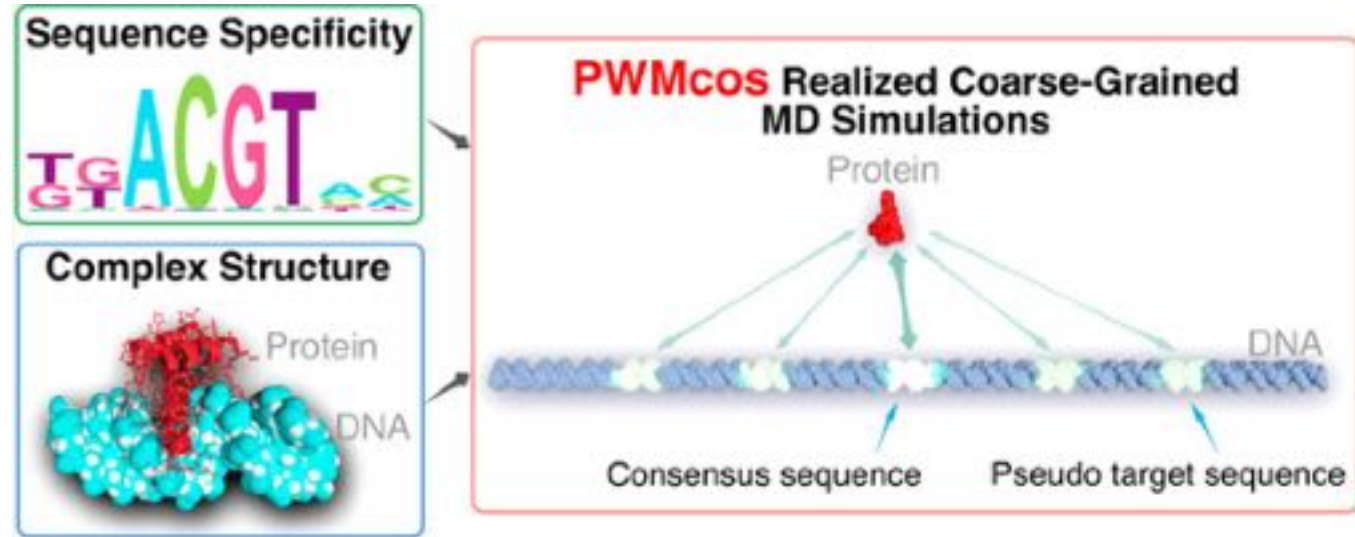


# Popular Coarse-grained Models for Biomolecules

## Protein-DNA:

### PWMcos

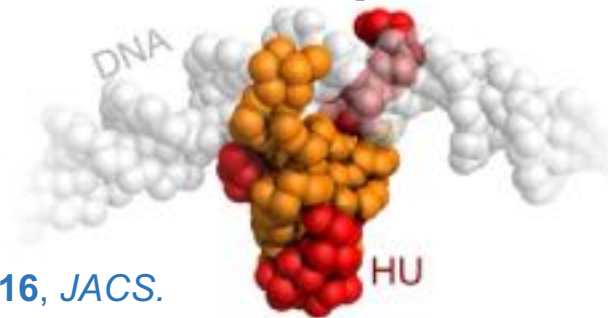
C.Tan, S. Takada, 2018, *JCTC*.



## Residue-level CG models: protein-DNA systems

~500 particles

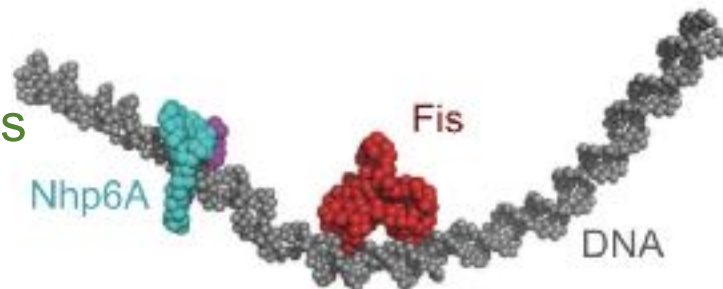
C.Tan, et al. 2016, *JACS*.



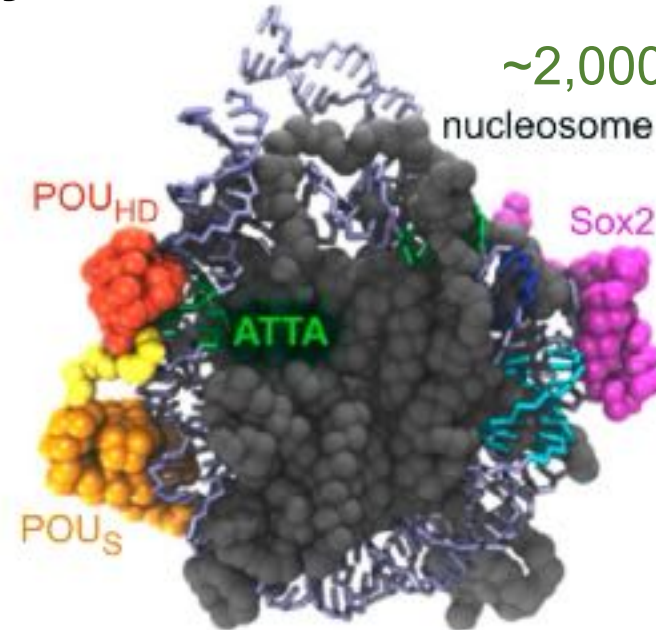
~1,000 particles



K. Kamagata et al. 2020, *NAR*; 2021, *NAR*.



~2,000 particles



Thanks to:  
Shoji Takada  
Giovanni Brandani  
@ Kyoto Univ.

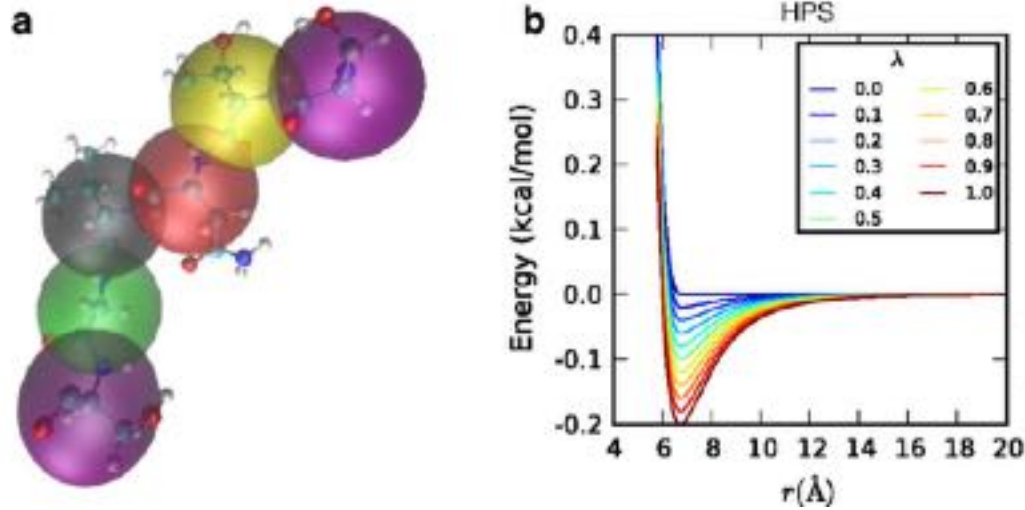
C.Tan, S. Takada, 2020, *PNAS*.



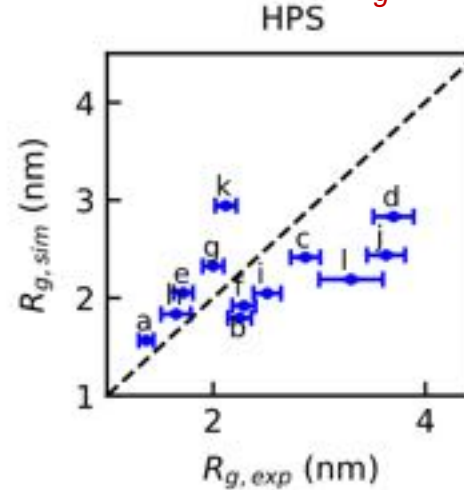
# HPS/KH Models for IDP

Dignon et al. 2018 PLoS Comput. Biol.

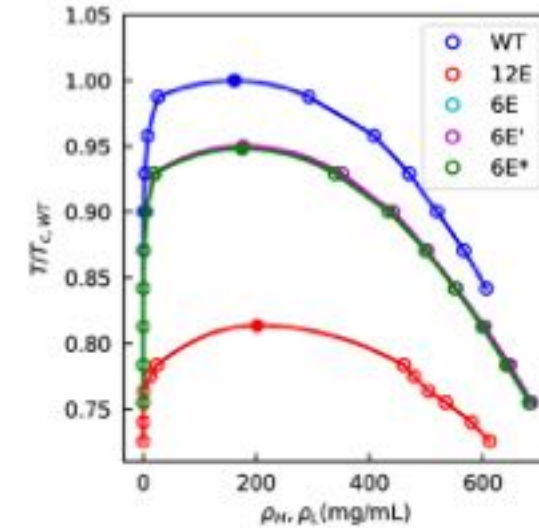
## Model resolution and potential function



## Calibration: $R_g$



## FUS WT and mutants



## Electrostatics

$$E_{ij}(r) = \frac{q_i q_j}{4\pi D r} \exp(-r/\kappa),$$

## Hydrophobicity scale (HPS) model

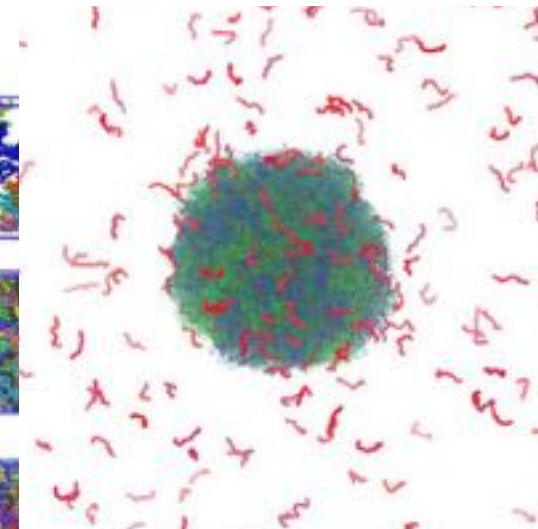
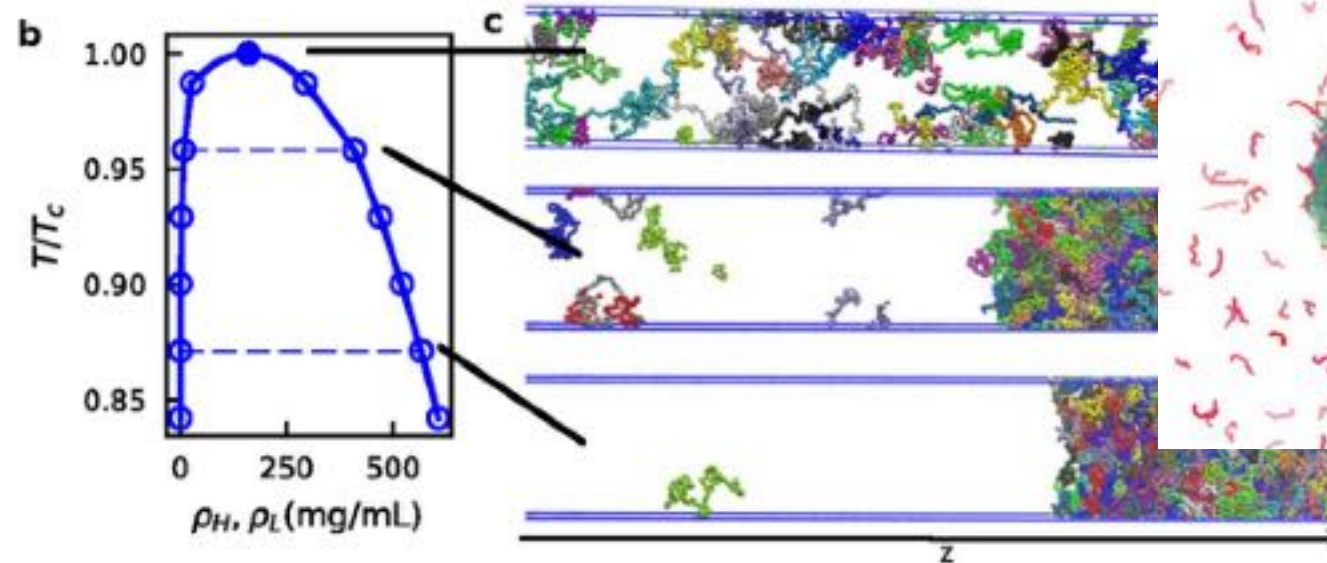
Ashbaugh-Hatch functional form

$$\Phi(r) = \begin{cases} \Phi_{LJ} + (1 - \lambda)\epsilon, & \text{if } r \leq 2^{1/6}\sigma \\ \lambda\Phi_{LJ}, & \text{otherwise} \end{cases}$$

Lennard-Jones potential

$$\Phi_{LJ} = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right].$$

## Slab method: simulating multi-chain systems



# Difficulties of Carrying out Coarse-grained MD

## “Barriers” of running CG MD simulations

### 1. To combine different models

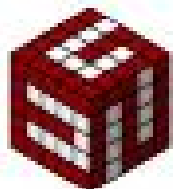
- Protein model: by Takada-group@Kyoto U.
- DNA model: by de Pablo-group@U. Chicago
- IDP model: by Mittal-group@Texas A&M U.
- ...

### 2. To apply to large-scale simulations

- Memory limit using *atomic decomposition*
- Computational efficiency
- Robustness and accuracy

## Part I: Implementation of CG Models

in



GENESIS

ATDYN

Generalized-ensemble simulation system

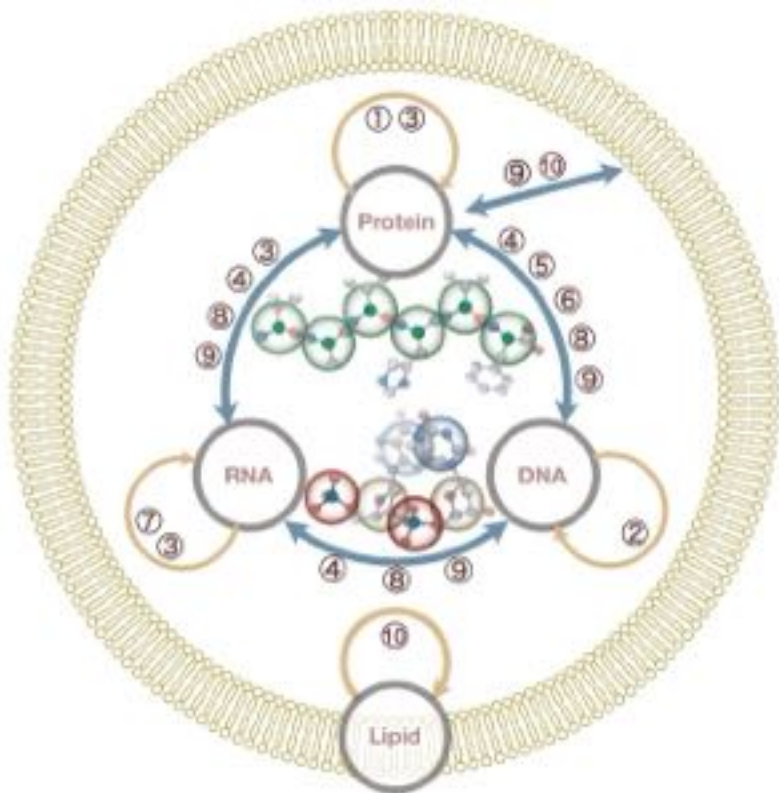
Jung et al. 2015 WIREs Comput. Mol. Sci.

Kobayashi et al. 2017 J. Comput. Chem.

# Implementation of CG Models in GENESIS

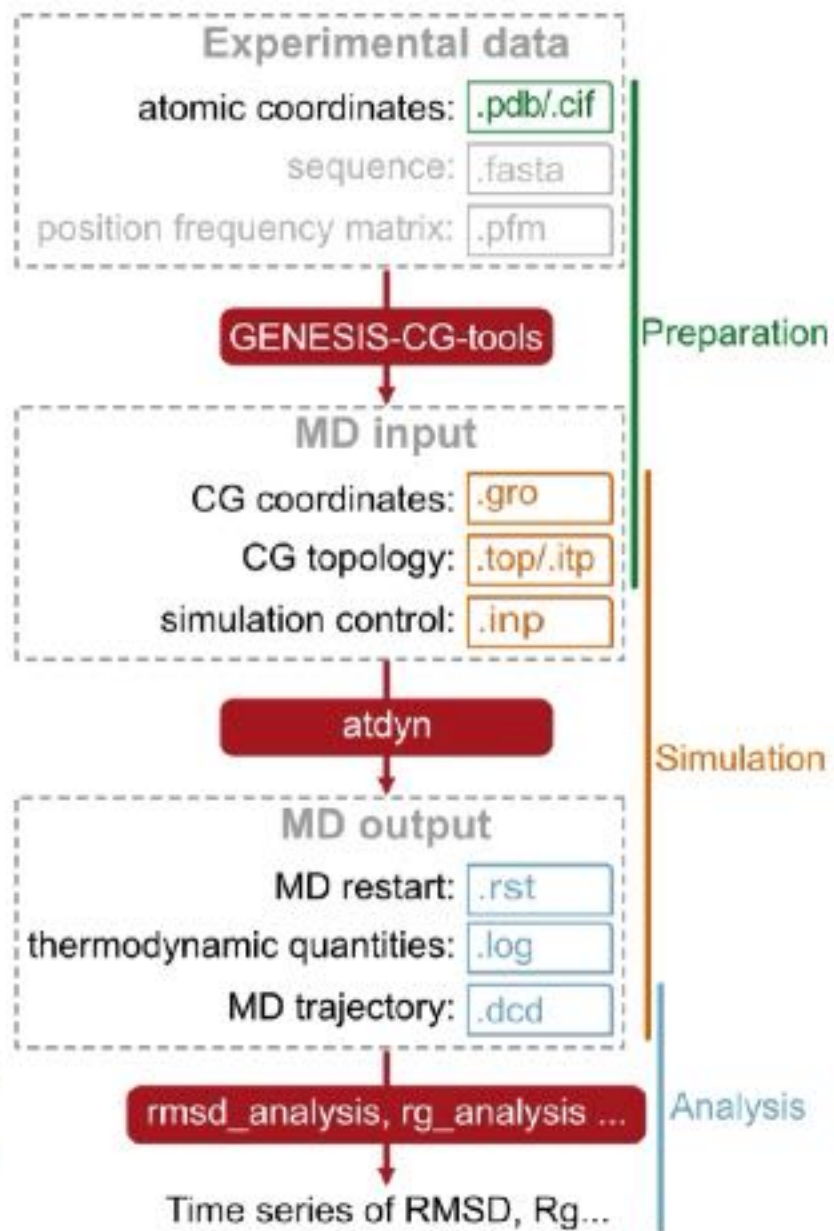
C.Tan et al. 2022, PLoS Comput. Biol.

## Interaction potentials

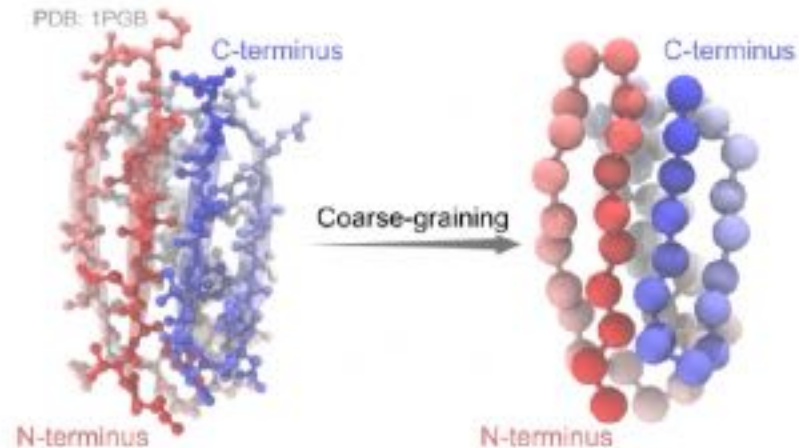


1. W. Li et al. (2014). *Proc. Natl. Acad. Sci.*
2. G.S. Freeman et al. (2014). *J. Chem. Phys.*
3. G.L. Dignon et al. (2018). *PLoS Comput. Biol.*
4. C. Clementi et al. (2000). *J. Mol. Biol.*
5. C. Tan & S. Takada (2018). *J. Chem. Theory Comput.*
6. G.B. Brandani et al. (2018) *Nucl. Acids Res.*
7. N. Hori. & S. Takada (2012). *J. Chem. Theory Comput.*
8. C. Tan & S. Takada (2016). *J. Am. Chem. Soc.*
9. P. Debye & E. Hückel (1923). *Physikalische Zeitschrift*
10. D Ugarte La Torre et al. (2023). *J. Chem. Phys.*

## Information flow of CG Sim.



## File preparation example



```
$ ls
PR01.pdb

# run GENESIS_CG_TOOL
$ genesis_cg_tool/src/aa_2_cg.jl PR01.pdb

$ ls
PR01.pdb PR01_cg.gro PR01_cg.top PR01_cg.itp
```

## GENESIS-cg-tool: (in Julia)

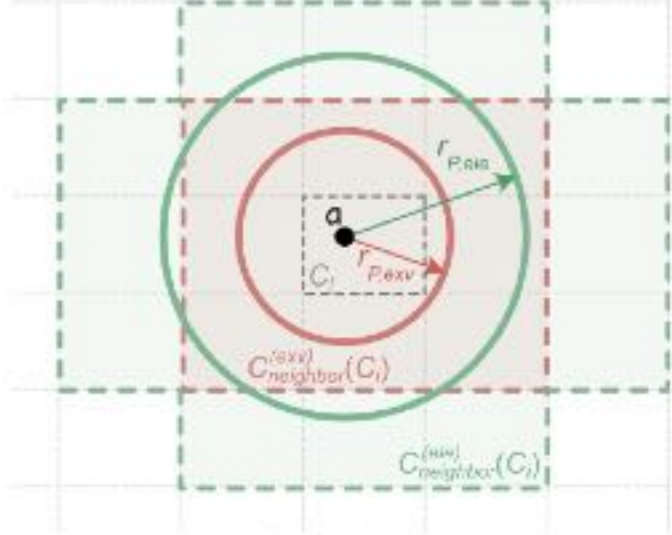
- [https://github.com/genesis-release-r-ocs/genesis\\_cg\\_tool](https://github.com/genesis-release-r-ocs/genesis_cg_tool)
- Online wiki:
  - [https://github.com/genesis-release-r-ocs/genesis\\_cg\\_tool/wiki](https://github.com/genesis-release-r-ocs/genesis_cg_tool/wiki)
- GENESIS source code + manual:
  - <https://www.r-ocs.riken.jp/labs/cbrt/genesis-version-2-0/>
- Tutorials:
  - <https://r-ocs.riken.jp/labs/cbrt/tutorials2022/>

0: Background | 1: GENESIS CG Models | 2: MD of Biomolecule LLPS

# Improve Computational Efficiency in CG Simulations

## Neighbor list design

Cell-linked lists  $O(N^2) \rightarrow O(N)$

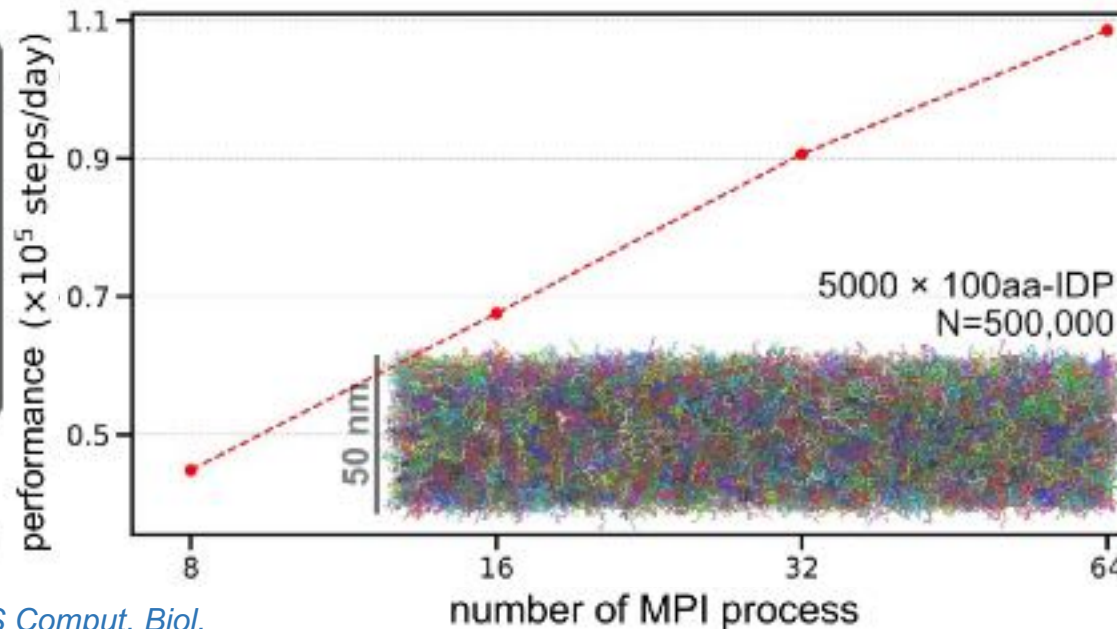
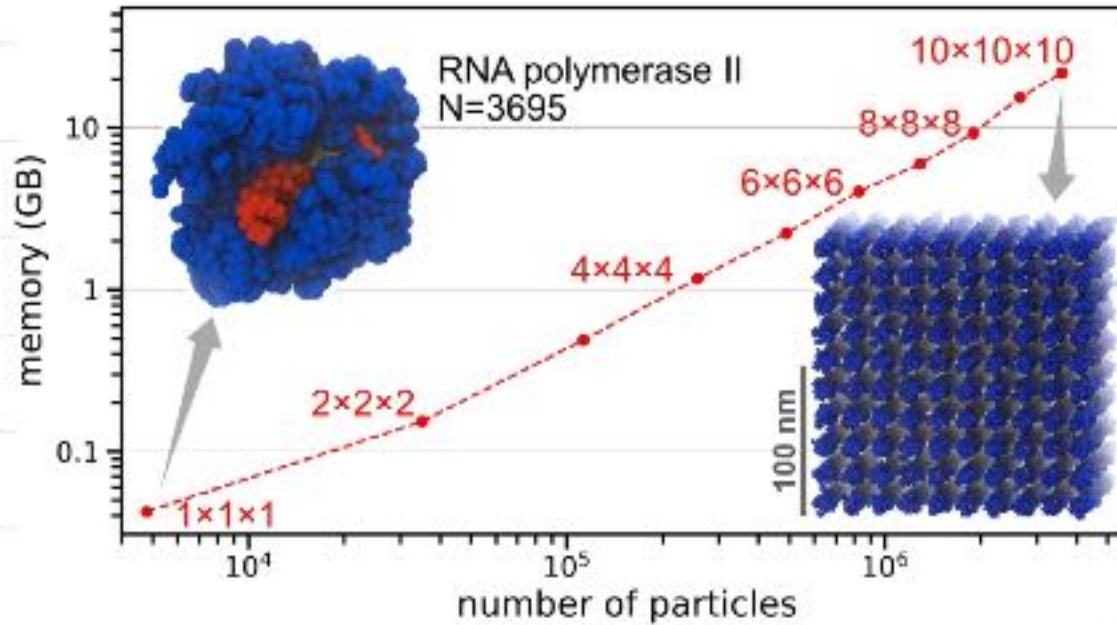


Pseudo code for assigning electrostatic pairlist:

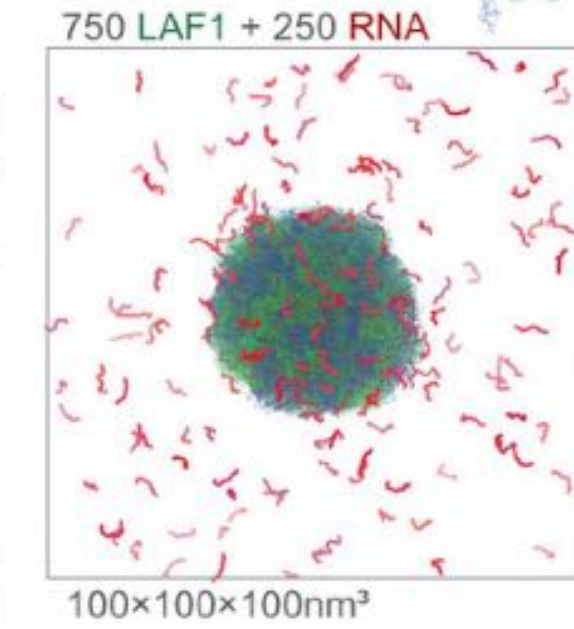
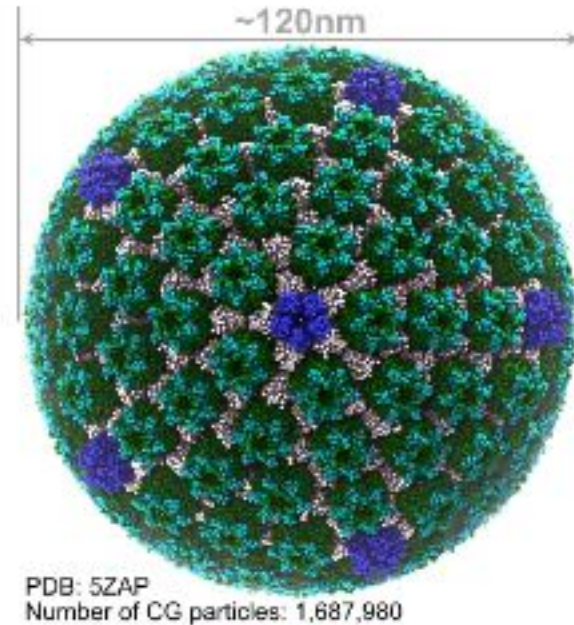
```

for a in all charged particles:
  C_i = the cell contains a
  for C_k in C_neighbor^{(ele)}(C_i):
    for b in L_ele(C_k):
      if r(a, b) <= r_{p,ele}:
        add b to a's neighborlist
    
```

## Memory and CPU benchmark



## Applications



Dr. Jaewoon Jung

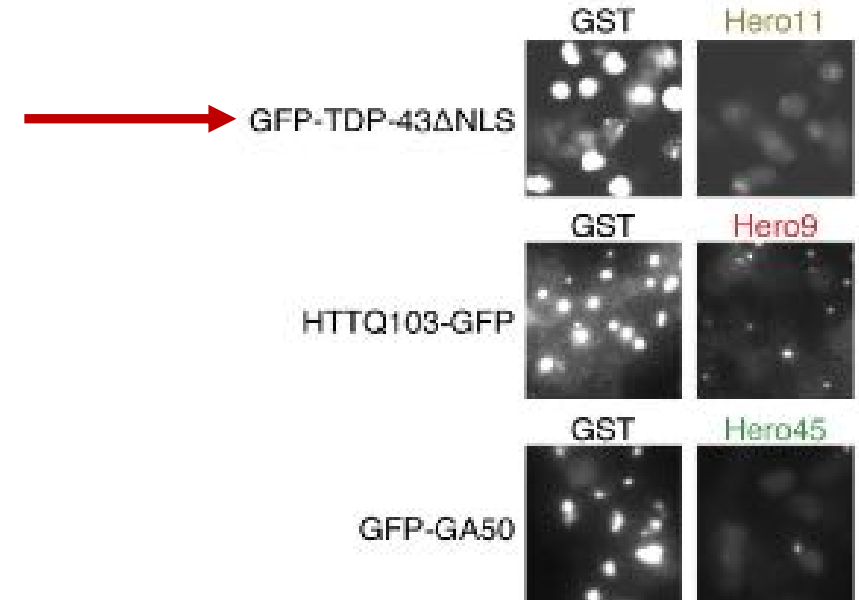
- Parallelization of random-number generation
- SIMD operations in non-bonded potentials

# Popular Coarse-grained Models for Biomolecules

## “Barriers” of running CG MD simulations

- ✓ 1. To combine different models
  - Protein model: by Takada-group@Kyoto U.
  - DNA model: by de Pablo-group@U. Chicago
  - IDP model: by Mittal-group@Texas A&M U.
  - ...
  
- ✓ 2. To apply to large-scale simulations
  - Memory limit for atomistic decomposition
  - Computational efficiency
  - Robustness and accuracy

Tsuboyama *et al.* 2020 *PLoS Biol.*



## Part I: Implementation of CG Models in GENESIS

## Part II: Application of GENESIS MD to HERO11 on supercomputers (Fugaku, Hokusai BW)

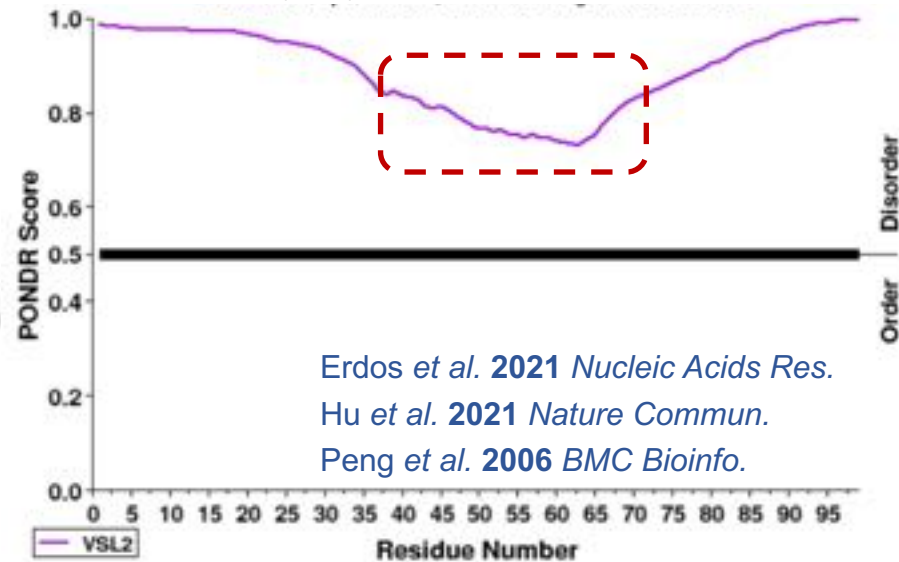
# Testify Hero11 $\alpha$ -helices Stability using Atomistic MD

Hero11-WT by AlphaFold2

Jumper *et al.* 2021 *Nature*

AF Database ID: Q9UNZ5

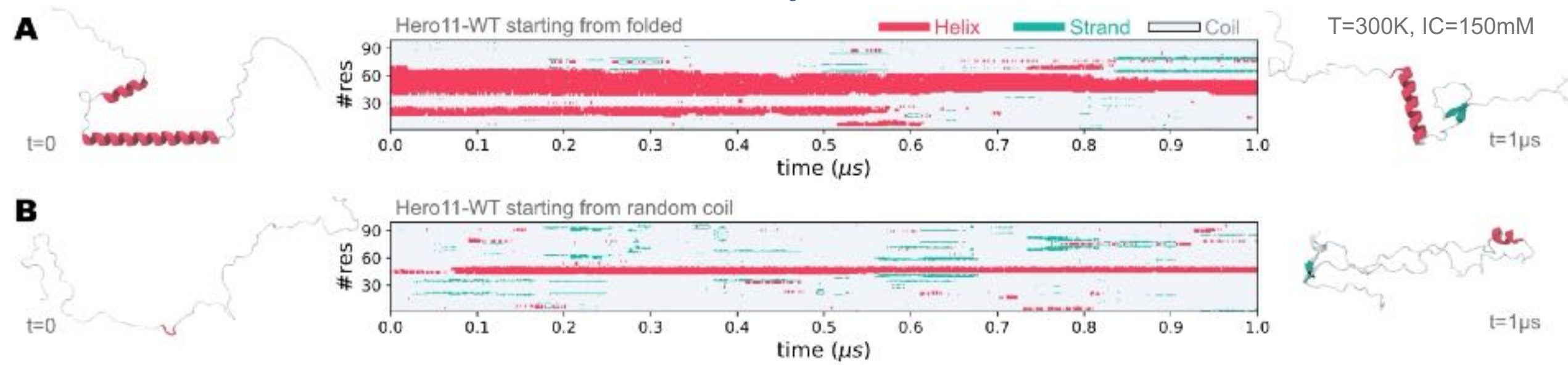
IDR propensity by IUPred3, fIDPnn, PONDR



Dr. Ai Niitsu

- AF-Predicted  $\alpha$ -helix can be unstable
- Convergence in atomistic MD takes long time

All-atom MD simulations: CHARM36m Force Field Huang *et al.* 2017 *Nature Methods*



# CG Modeling of Hero11 and TDP-43

C.Tan, A. Niitsu, Y. Sugita, *JACS Au*, 2023.

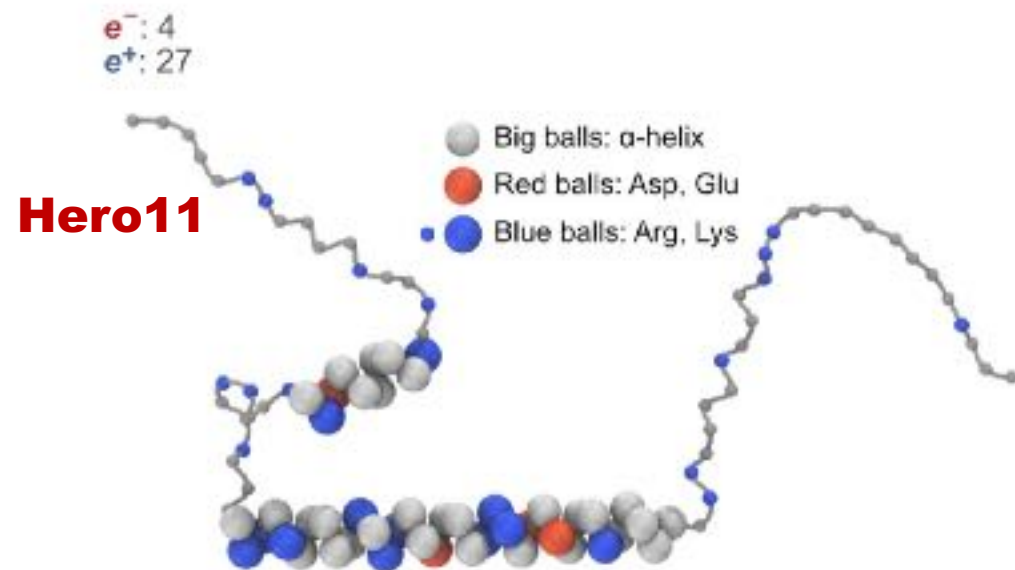
0: Background

1: GENESIS CG Models

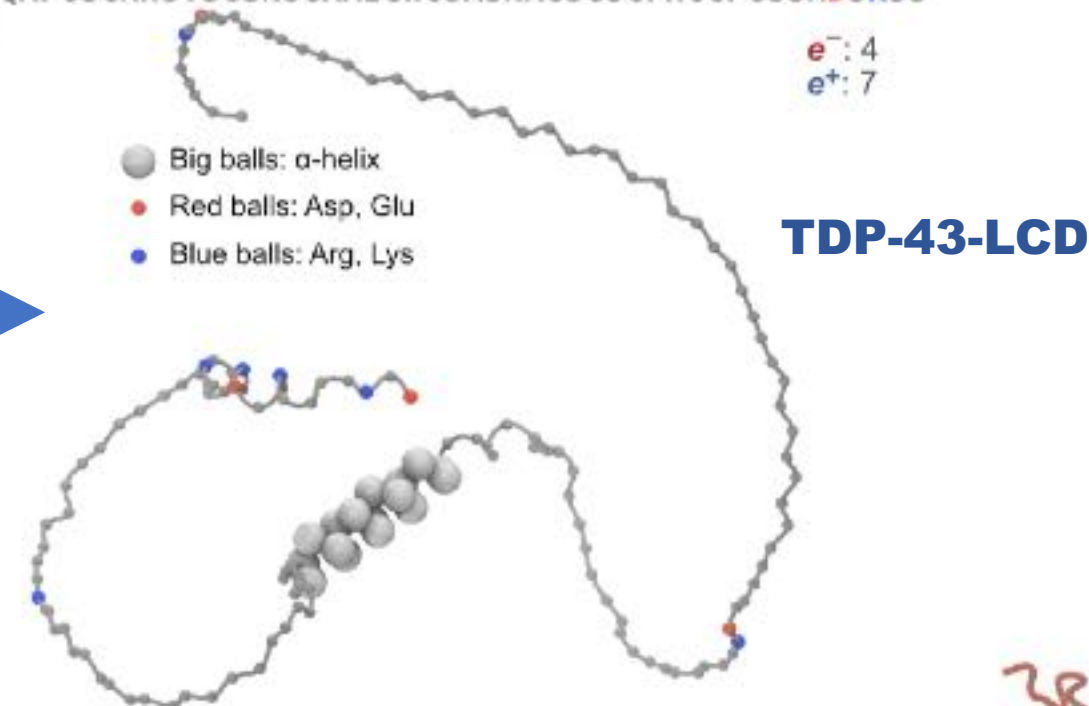
2: MD of Biomolecule LLPS

MAQGQRKFQAHKPAKSKTAAAAASEKNRGPRKGGRVVIAPKKARVVQQQKLK  
KNLEVGIRKKIEHDVVMKASSSLPKKLALLKAPAKKKGAAAATSSKTPS

EPKHNSNRQLERSGRFGGNPGGFGNQGGFGNSRGGGAGLGNNQGSNMGGG  
MNFGAFSINPAMMAAAQAAALOSSWGMGMLASQQNQSGPSGNNQNGNMQ  
REPNQAFGSGNNSYSGSNSGAAIGWGSASNAGSGSGFNNGGFGSSMDKSS  
GWGM



regulate



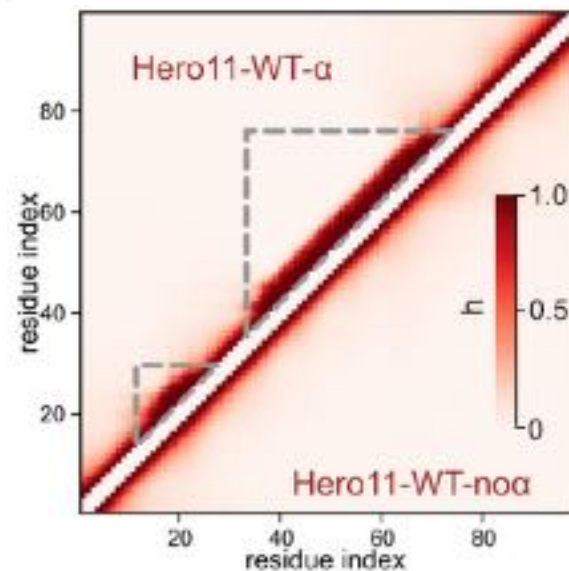
## CG Models:

- IDR: **HPS**  
(main contributors for LLPS)

*Dignon et al. 2018 PLoS Comput Biol.*

- folded domains: **AICG2+**  
(maintain  $\alpha$ -helical structure)

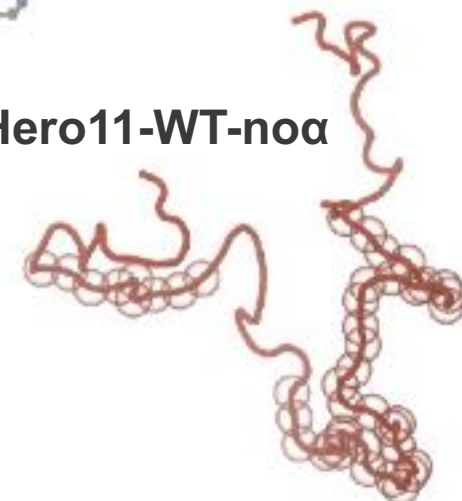
*Li et al. 2014 PNAS*



Hero11-WT- $\alpha$



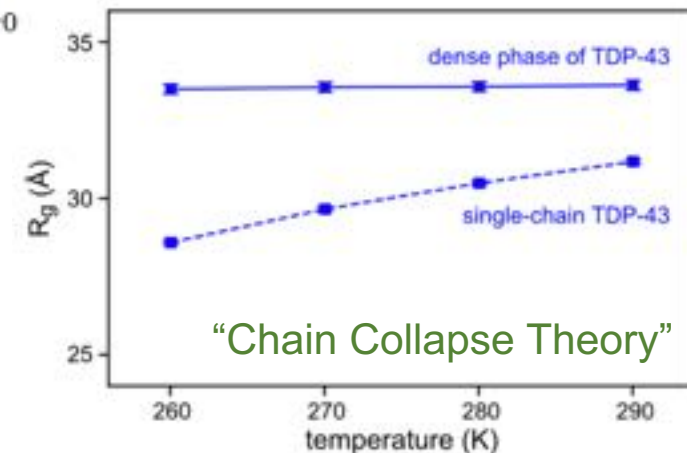
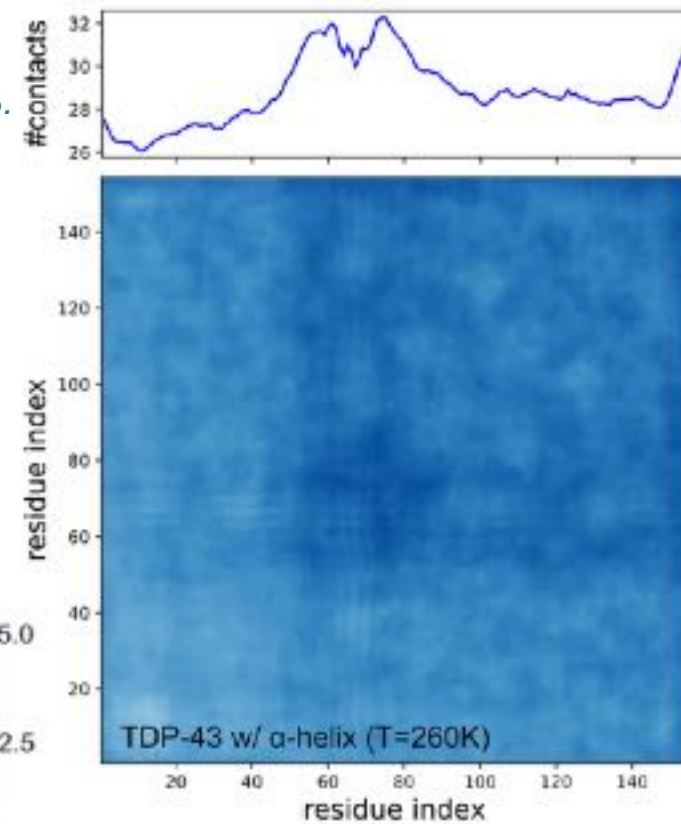
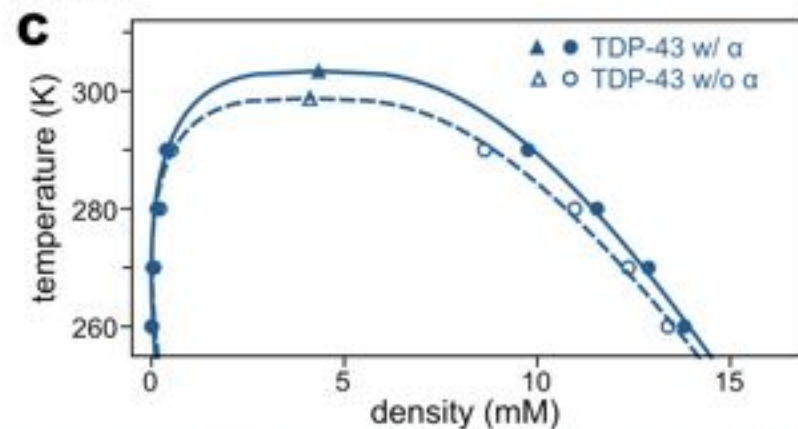
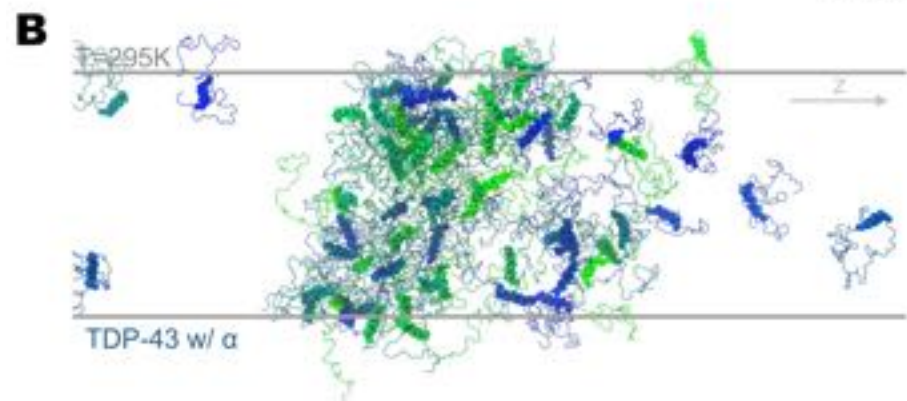
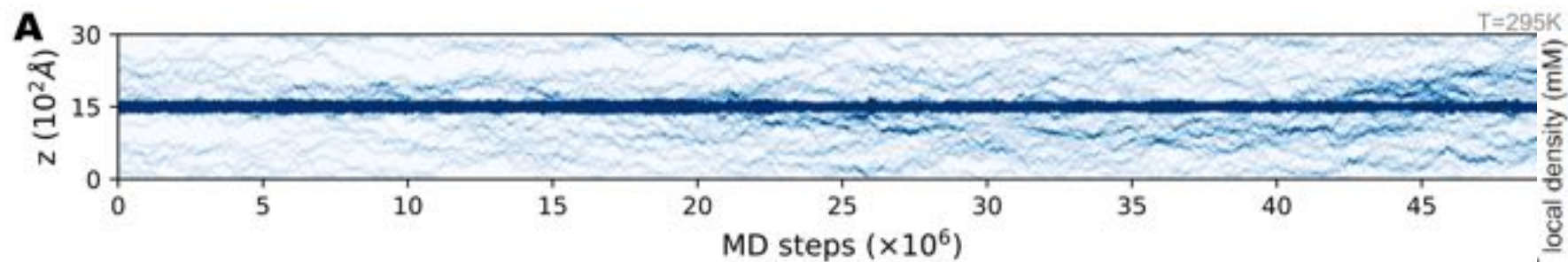
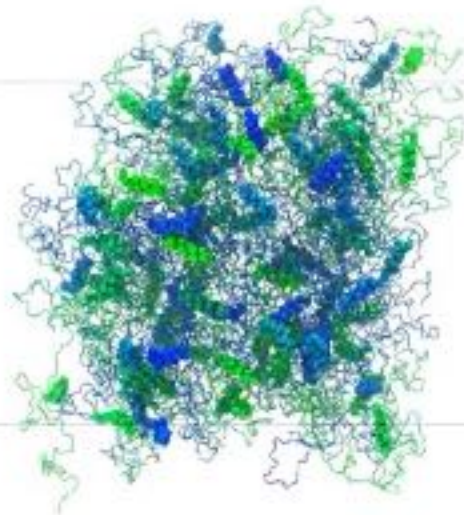
Hero11-WT-no $\alpha$



# Homotypic Condensation of TDP-43

C.Tan, A. Niitsu, Y. Sugita, *JACS Au*, 2023.

- 100 TDP-43
- $180 \times 180 \times 3000 \text{ \AA}^3$
- $T=250 \sim 350 \text{ K}$
- $IC=150 \text{ mM}$



“Chain Collapse Theory”

Consistent with Conicella et al. 2020 PNAS

Chou et al. 2006 J. Phys. Chem. Lett.



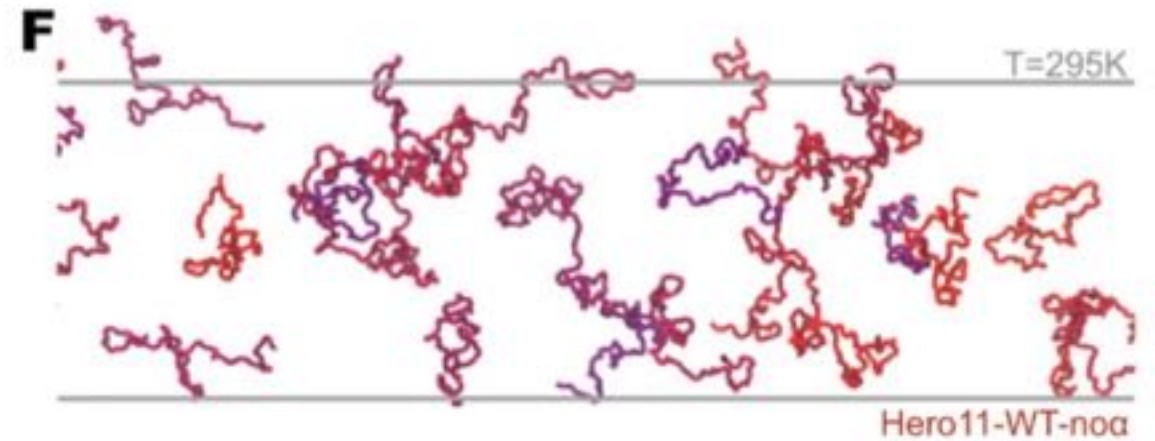
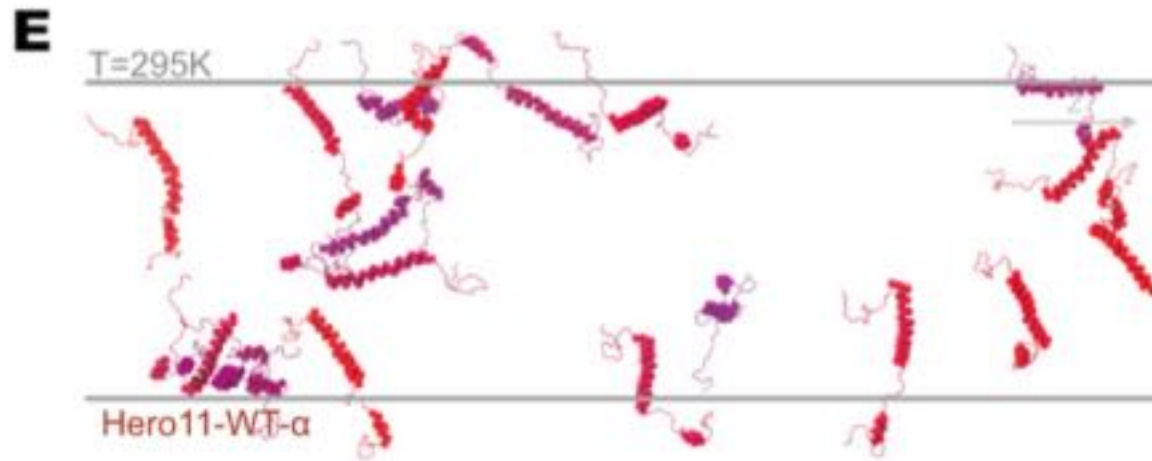
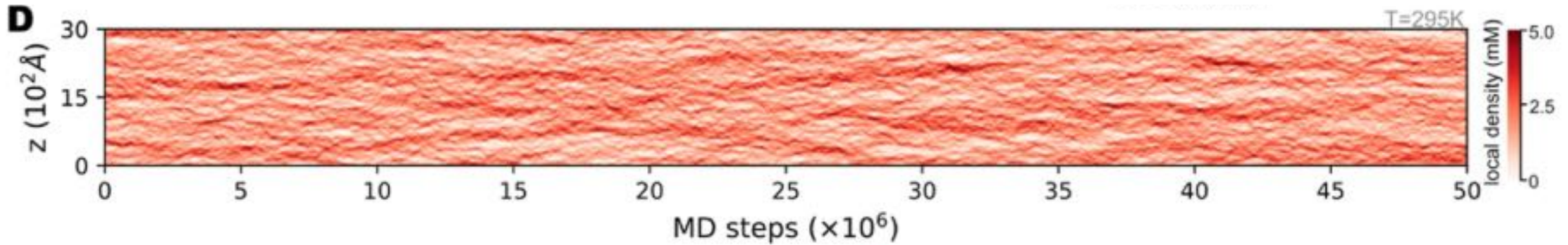
# Homotypic Condensation of Hero11 WT- $\alpha$ and WT-no $\alpha$

- 100 Hero11-WT- $\alpha$
- $180 \times 180 \times 2000 \text{ \AA}^3$
- $T=50 \sim 350 \text{ K}$
- $IC=150 \text{ mM}$

$T_c \ll 295 \text{ K}$

**Hero11: low critical temperature**

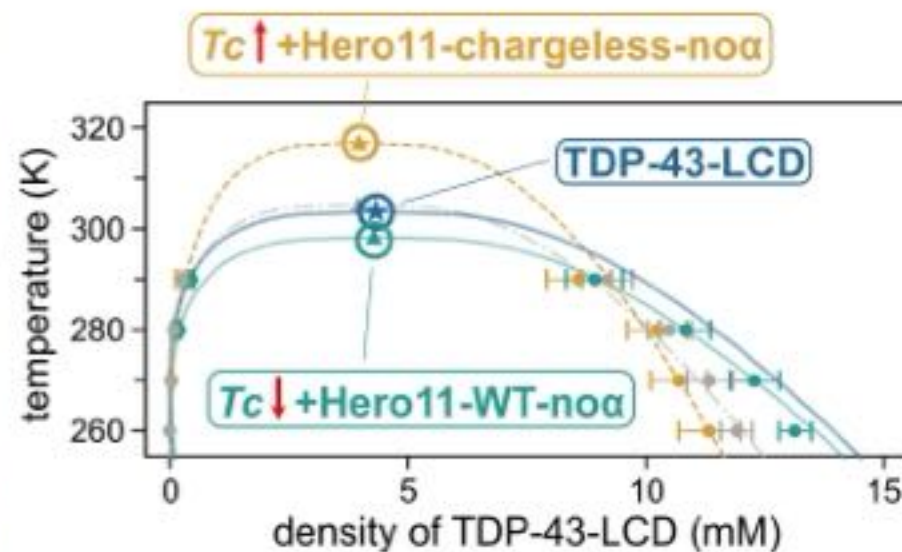
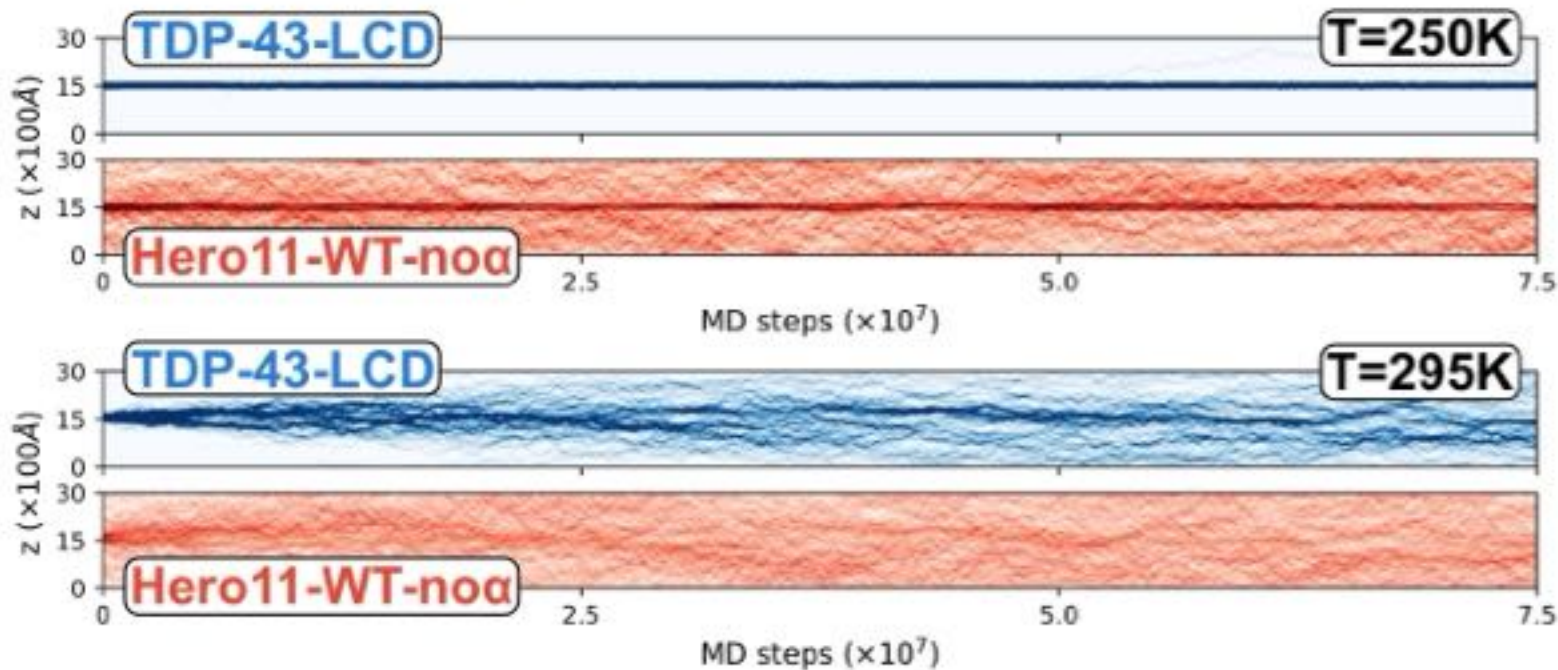
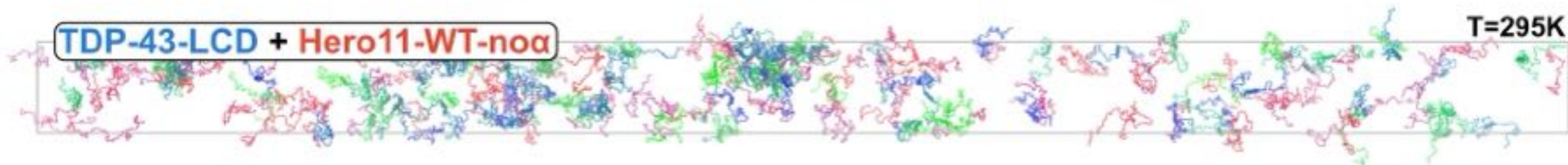
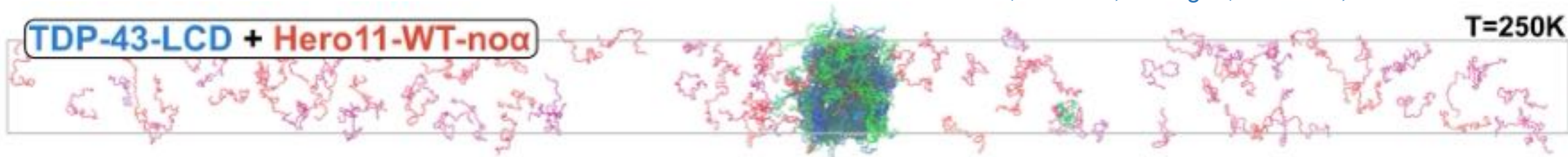
C.Tan, A. Niitsu, Y. Sugita, *JACS Au*, 2023.



# Heterotypic Condensation of TDP-43 + Hero11-no- $\alpha$

C.Tan, A. Niitsu, Y. Sugita, *JACS Au*, 2023.

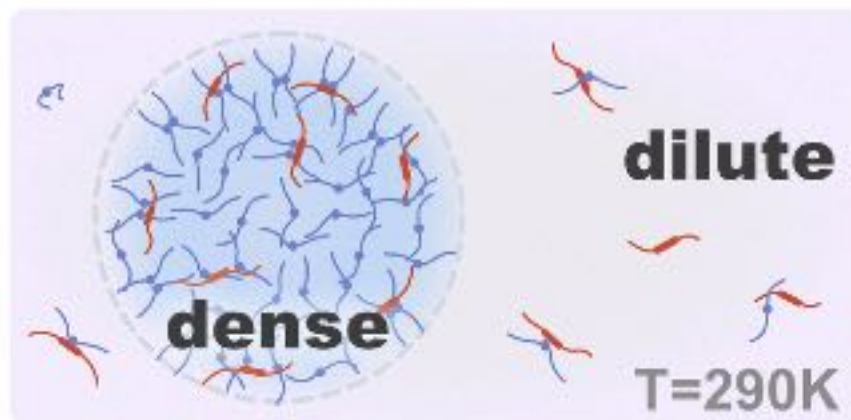
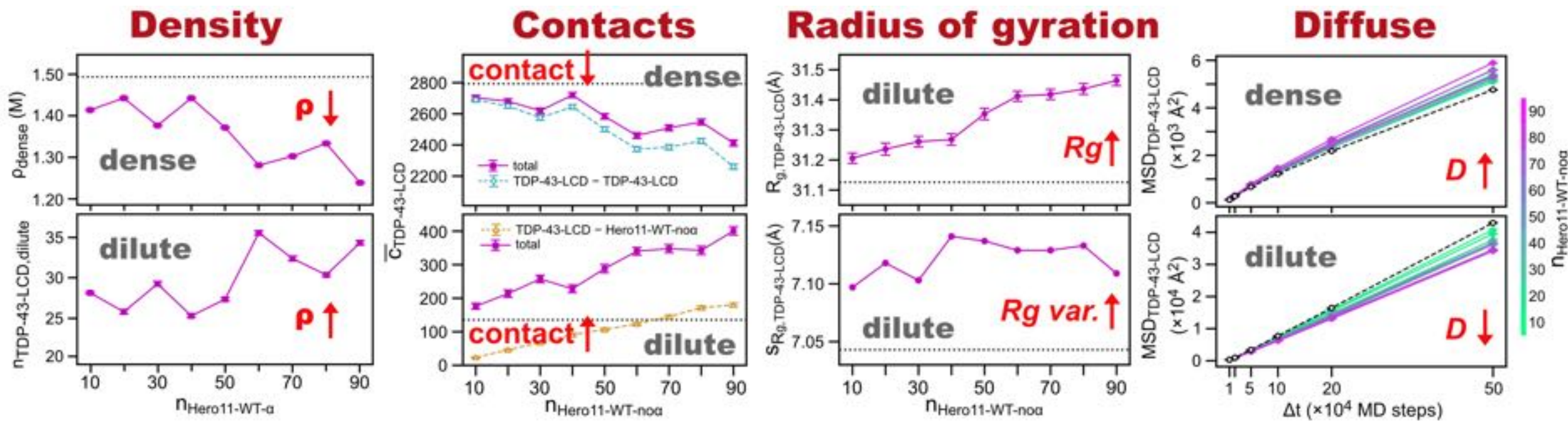
0: Background | 1: GENESIS CG Models | 2: MD of Biomolecule LLPS



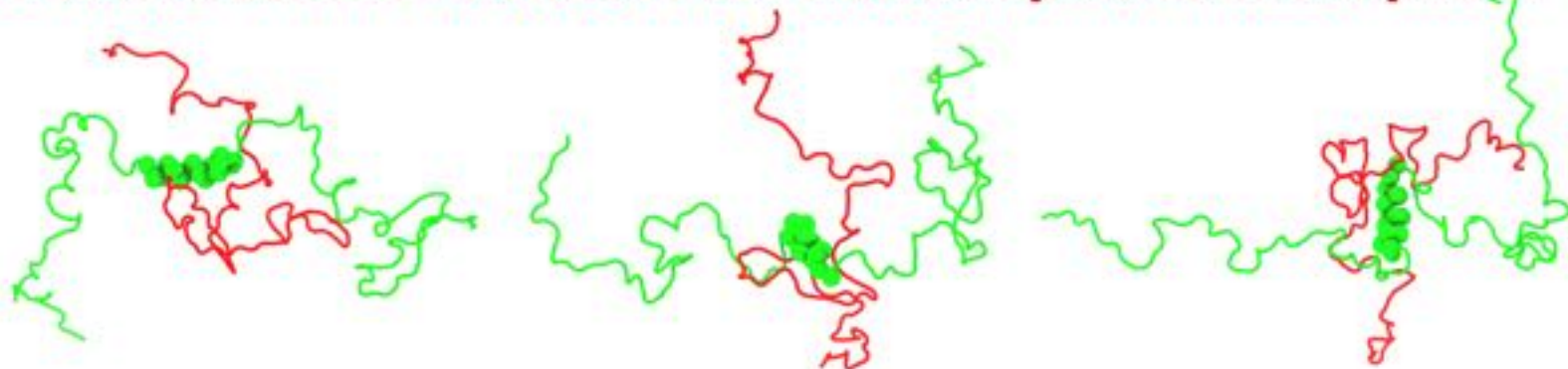
**Hero11 lowers  $T_c$  of TDP-43**

# Structure, dynamics, interaction of Hero11-no- $\alpha$ and TDP-43 in two phases

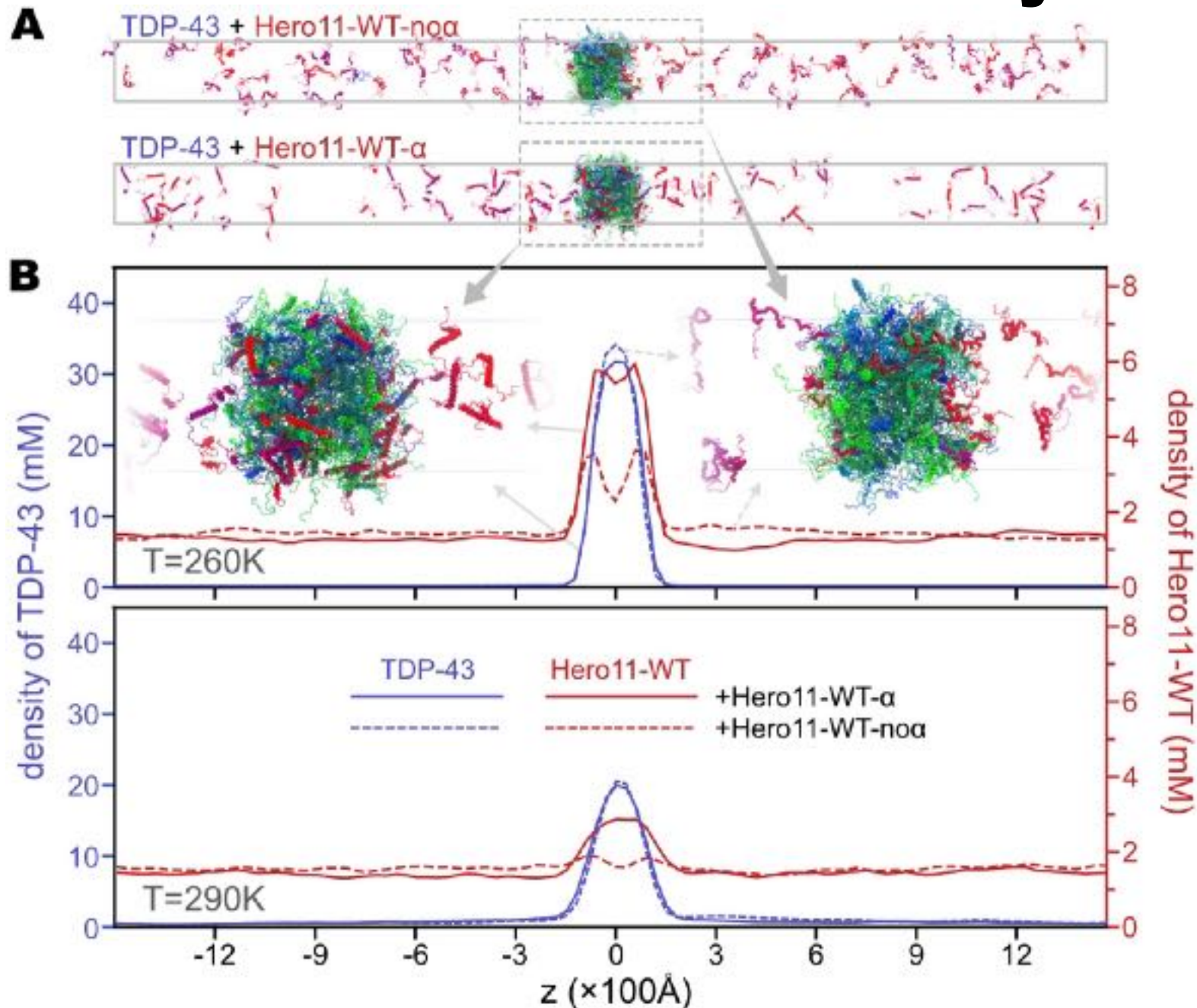
C.Tan, A. Niitsu, Y. Sugita, *JACS Au*, 2023.



## Structure of TDP-43/Hero11-WT-no $\alpha$ complex in dilute phase



# Effects of $\alpha$ -helical secondary structure in Hero11



C.Tan, A. Niitsu, Y. Sugita, *JACS Au*, 2023.

Distribution of Hero11:

- $\alpha$ -helical model:  
interior of condensate
- IDP model:  
surface of condensate

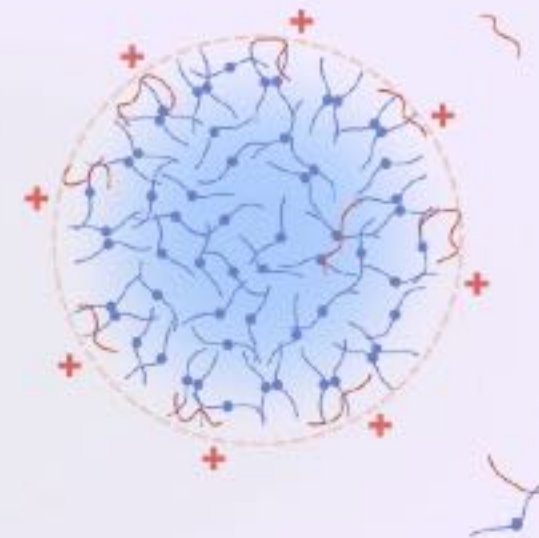
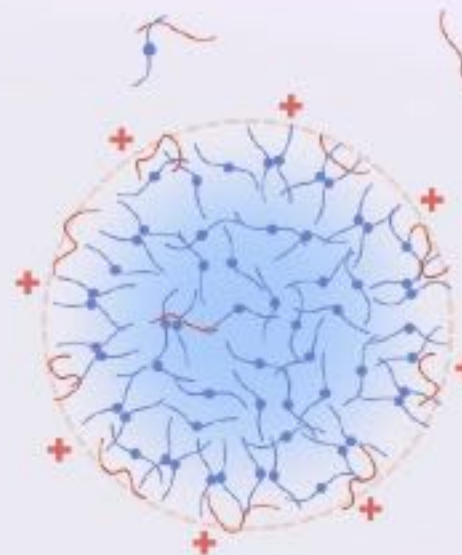
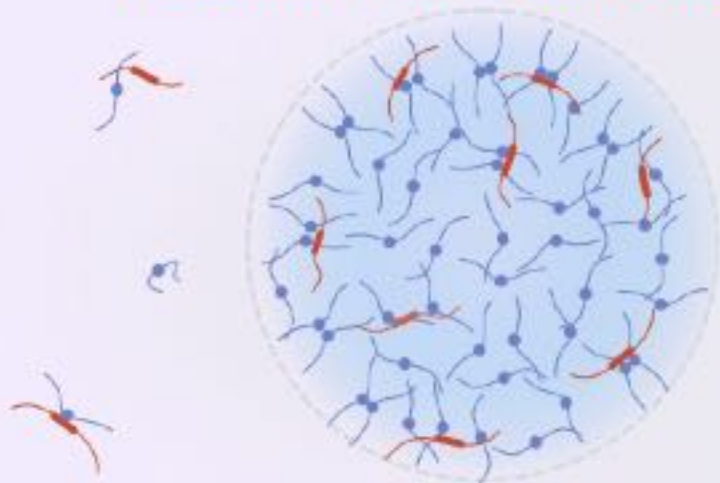
Surface electrostatics & fusion propensity

Welsh et al. 2022 *Nano Lett.*

# Summary: Hero11's LLPS-Regulation Mechanisms

C.Tan, A. Niitsu, Y. Sugita, *JACS Au*, 2023.

① TDP-43-LCD changes in dense phase:  
reduced inter-chain contacts, faster diffusion



② TDP-43-LCD changes in dilute phase:  
increased inter-chain contacts, larger variance of  $R_g$

③ possible effects of surface charge:  
decelerated droplet fusion?

Welsh *et al.* 2022 *Nano Lett.*

**Remaining Question:**

**How to verify mechanism no. 3?**

- We need a more powerful tool!

# Improve Computational Efficiency in CG Simulations



## GENESIS

Generalized-ensemble simulation system

### ATDYN

Atomic decomposition dynamics

### SPDYN

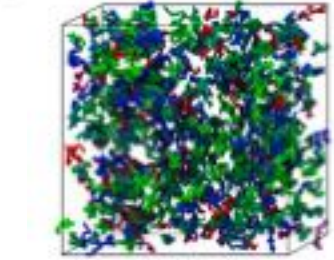
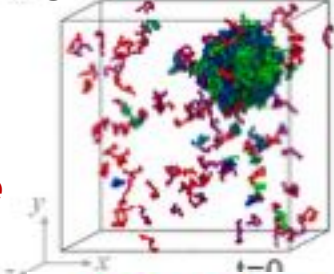
Spatial decomposition dynamics

### CGDYN

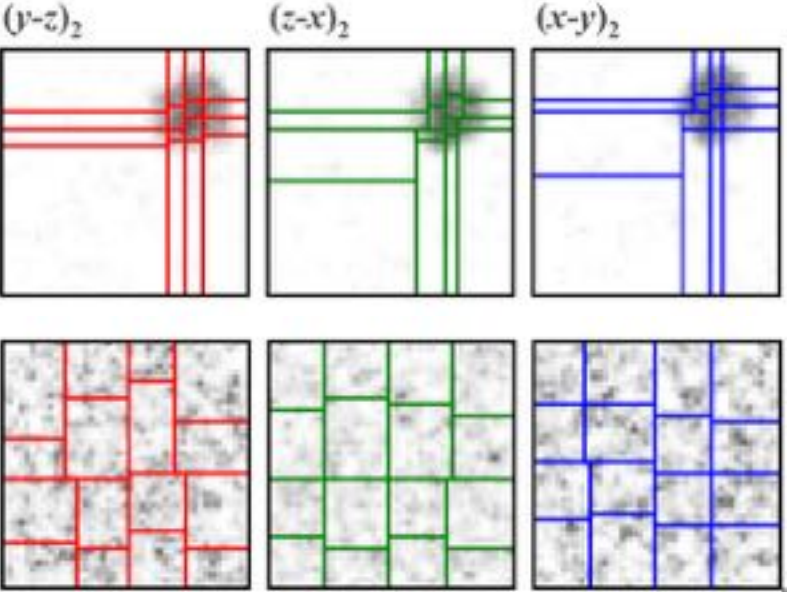
CG molecular dynamics

*Dynamic domain decomposition scheme to address the nonuniform distribution*

#### CGDYN



$t=9 \times 10^6$  steps



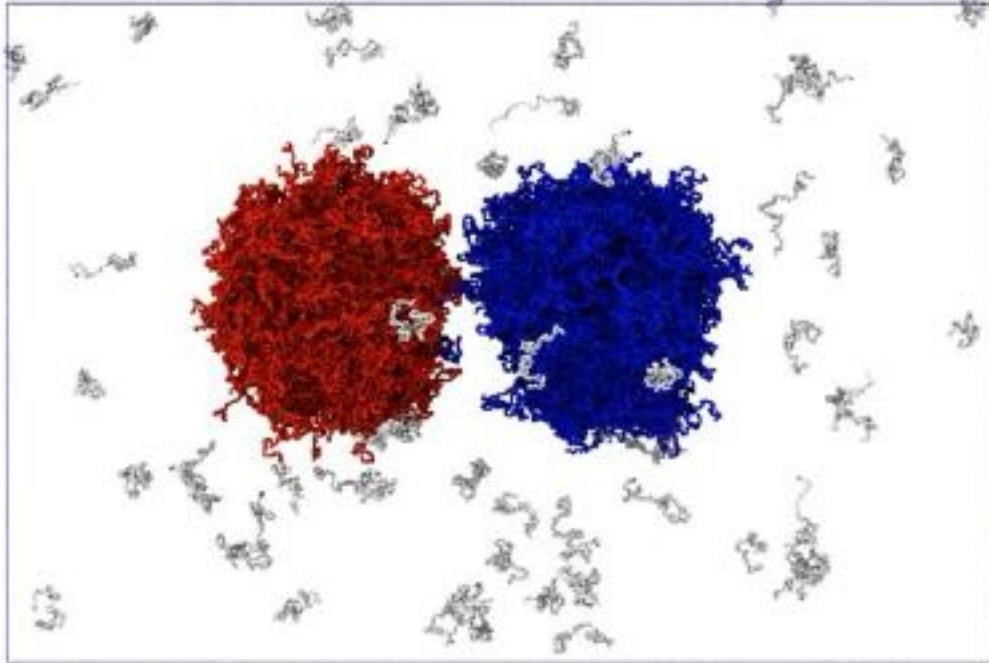
Dr. Jaewoon Jung

Jung, Tan, & Sugita, *bioRxiv*, 2023

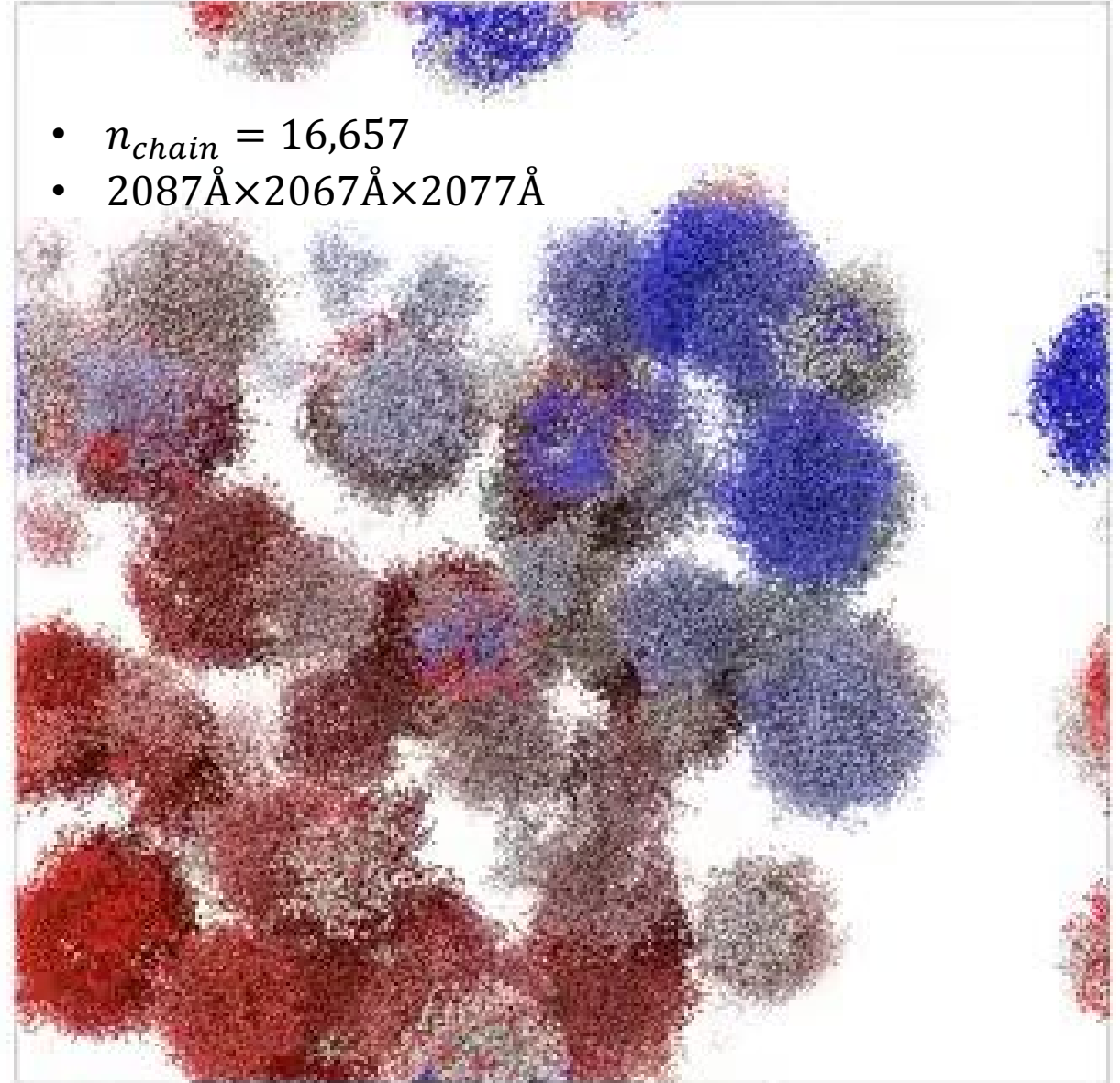
# CG MD of Droplets Going to Ultra-Large Scales

Component: TDP-43-LCD

- $n_{chain} = 1,000$
- $1000\text{\AA} \times 1000\text{\AA} \times 1500\text{\AA}$



- $n_{chain} = 16,657$
- $2087\text{\AA} \times 2067\text{\AA} \times 2077\text{\AA}$

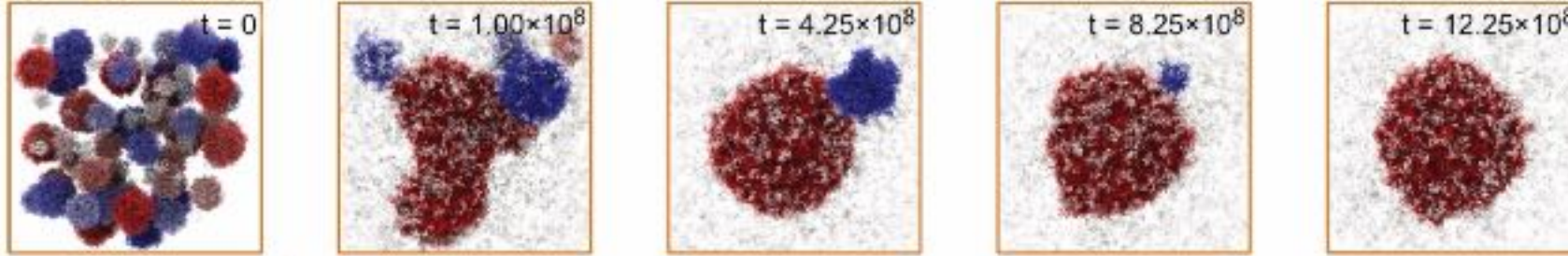


Jung, Tan, & Sugita, *bioRxiv*, 2023

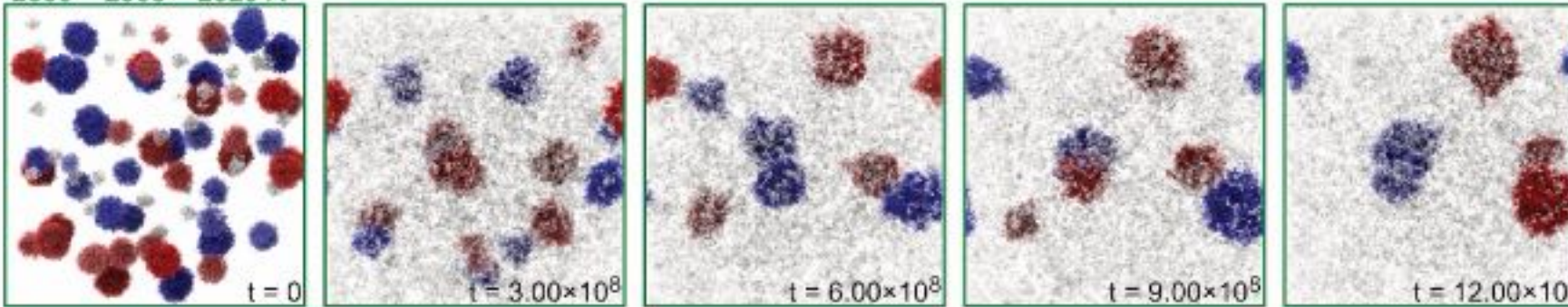
**Simulations carried out on Fugaku**

# CG MD of Droplets Going to Ultra-Large Scales

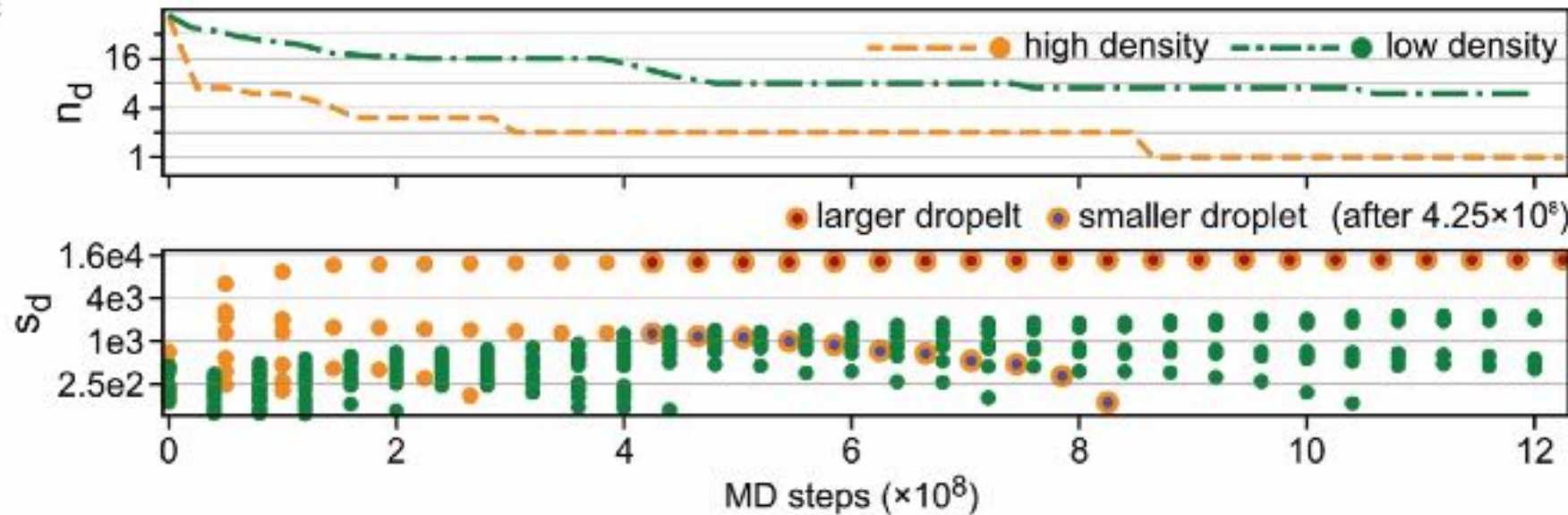
**a**  $2087 \times 2067 \times 2077 \text{ \AA}^3$



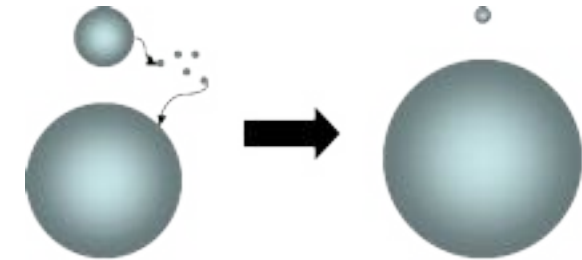
**b**  $2899 \times 2903 \times 2929 \text{ \AA}^3$



**c**

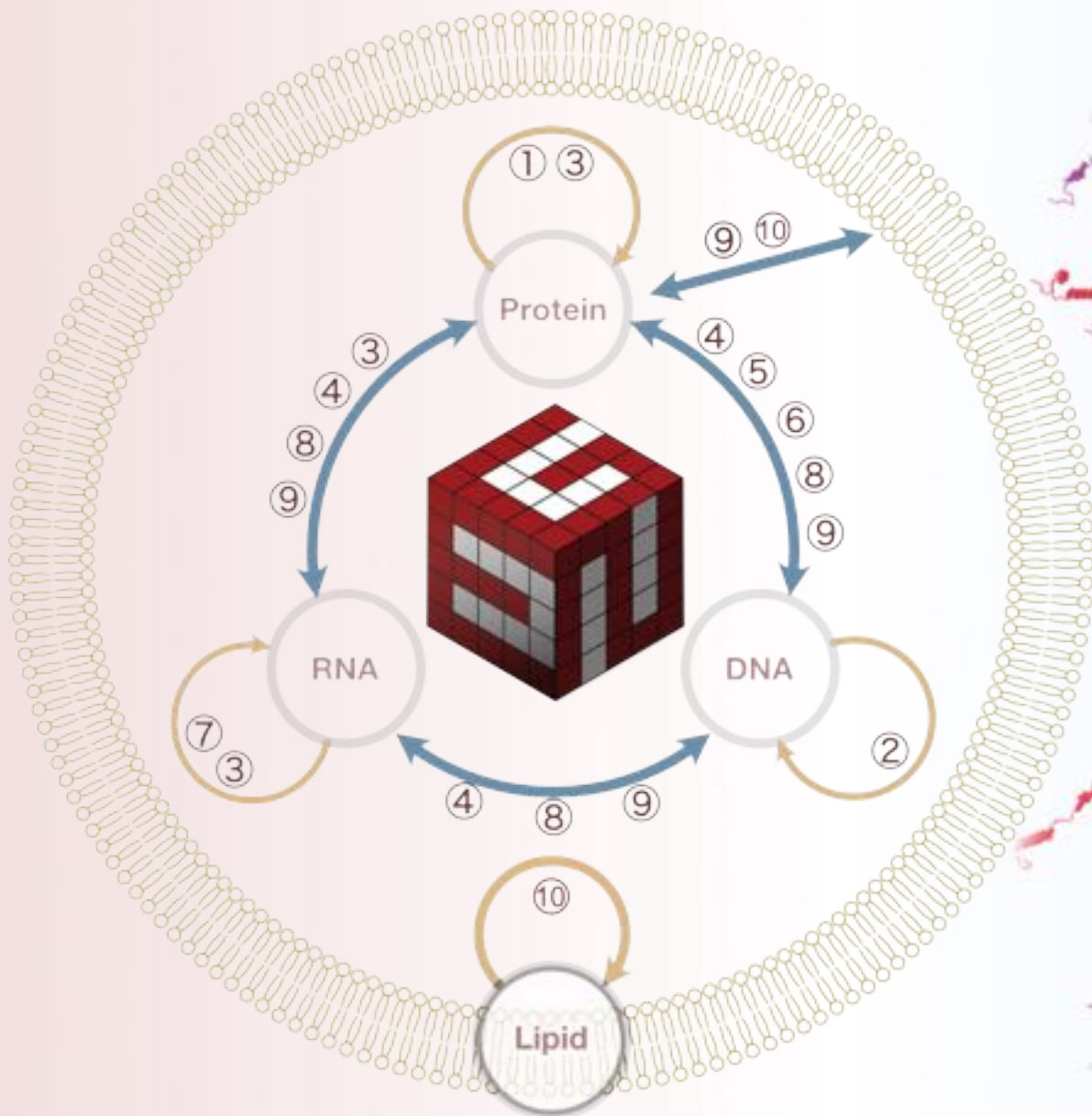


The first direct observation of  
"Ostwald ripening"

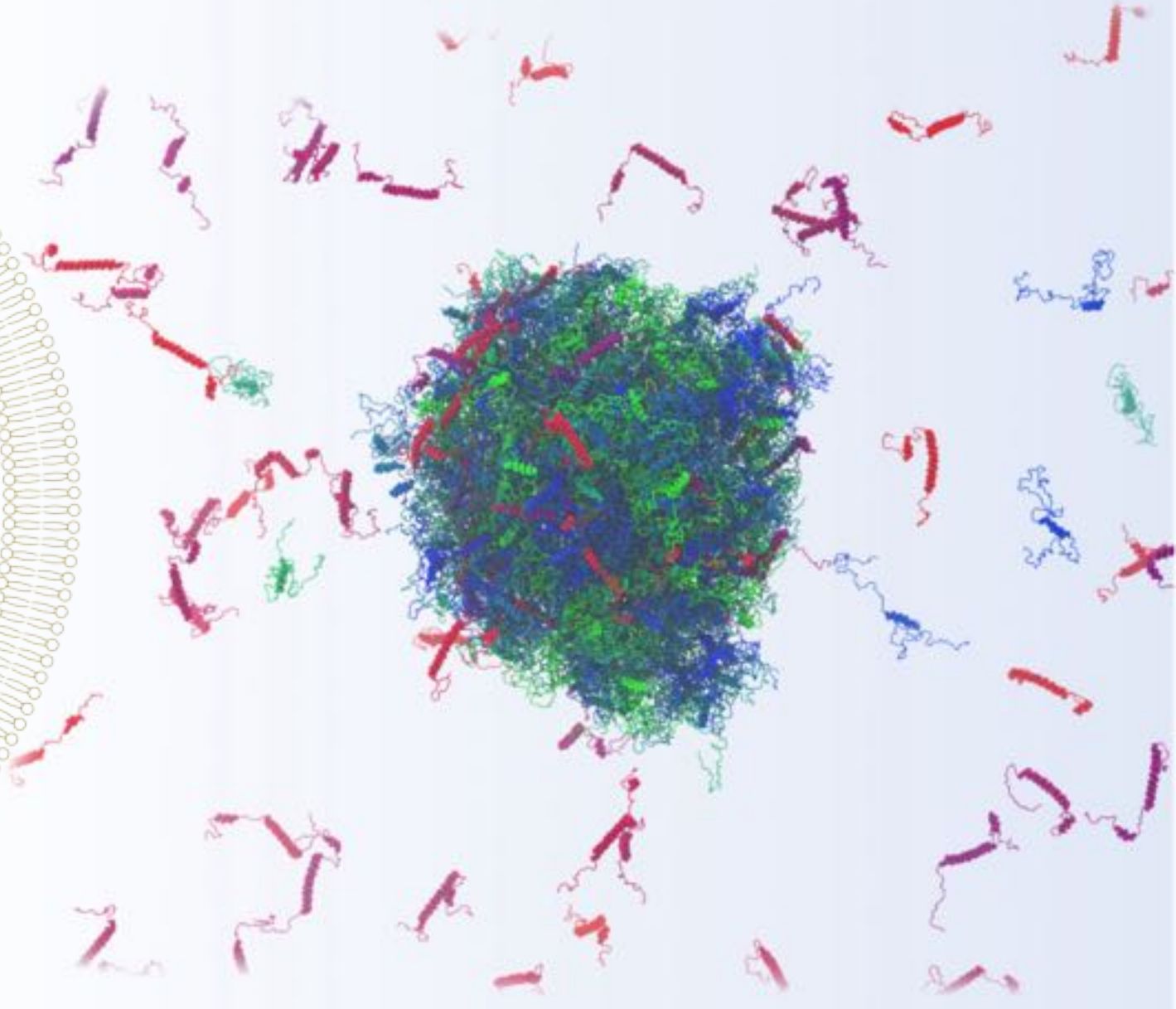




- **Developments of CG in GENESIS**  
User-friendly high-performance simulation tool



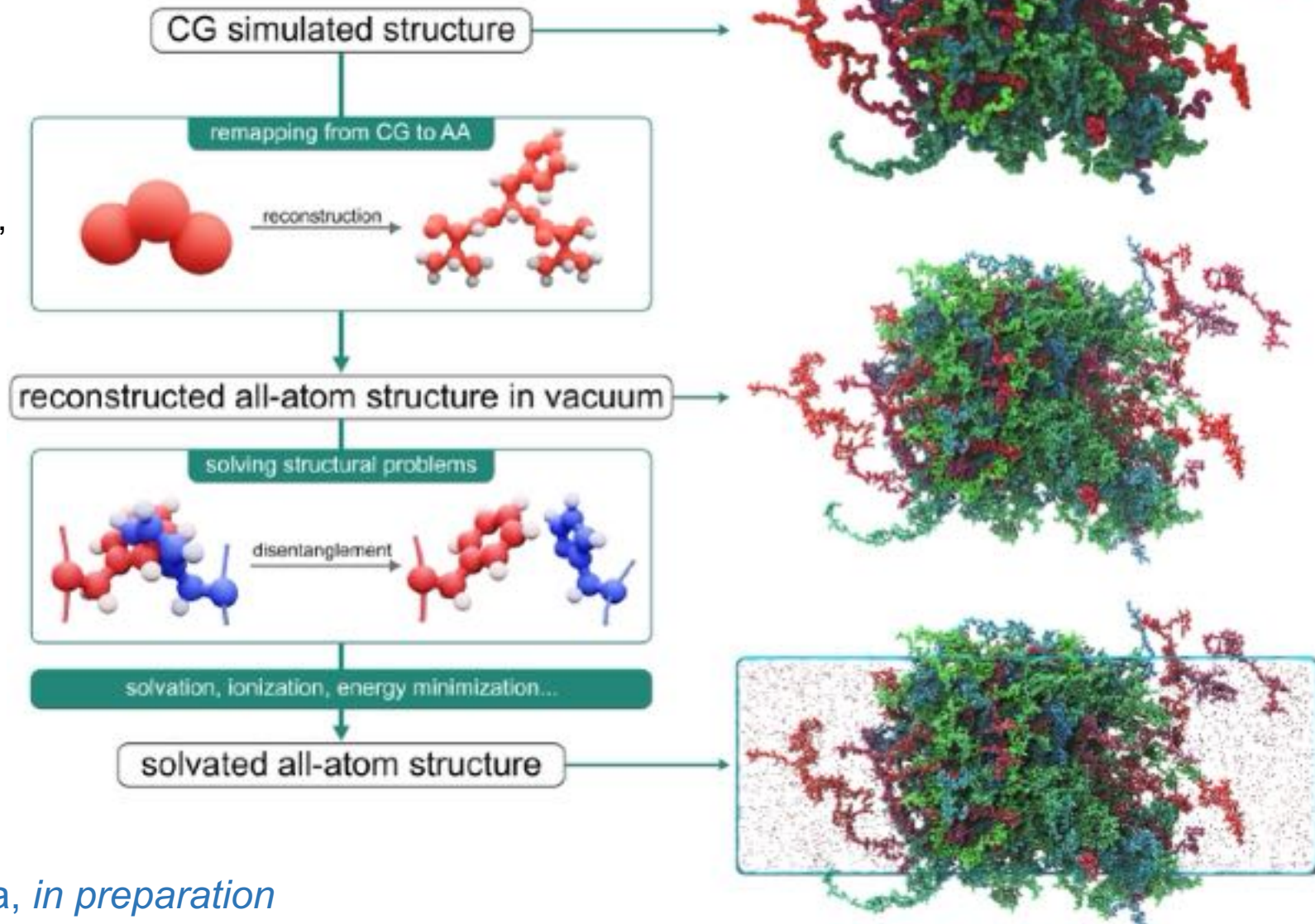
- **Regulation of biomolecular condensation**  
High-charged IDPs and their interactions regulate LLPS



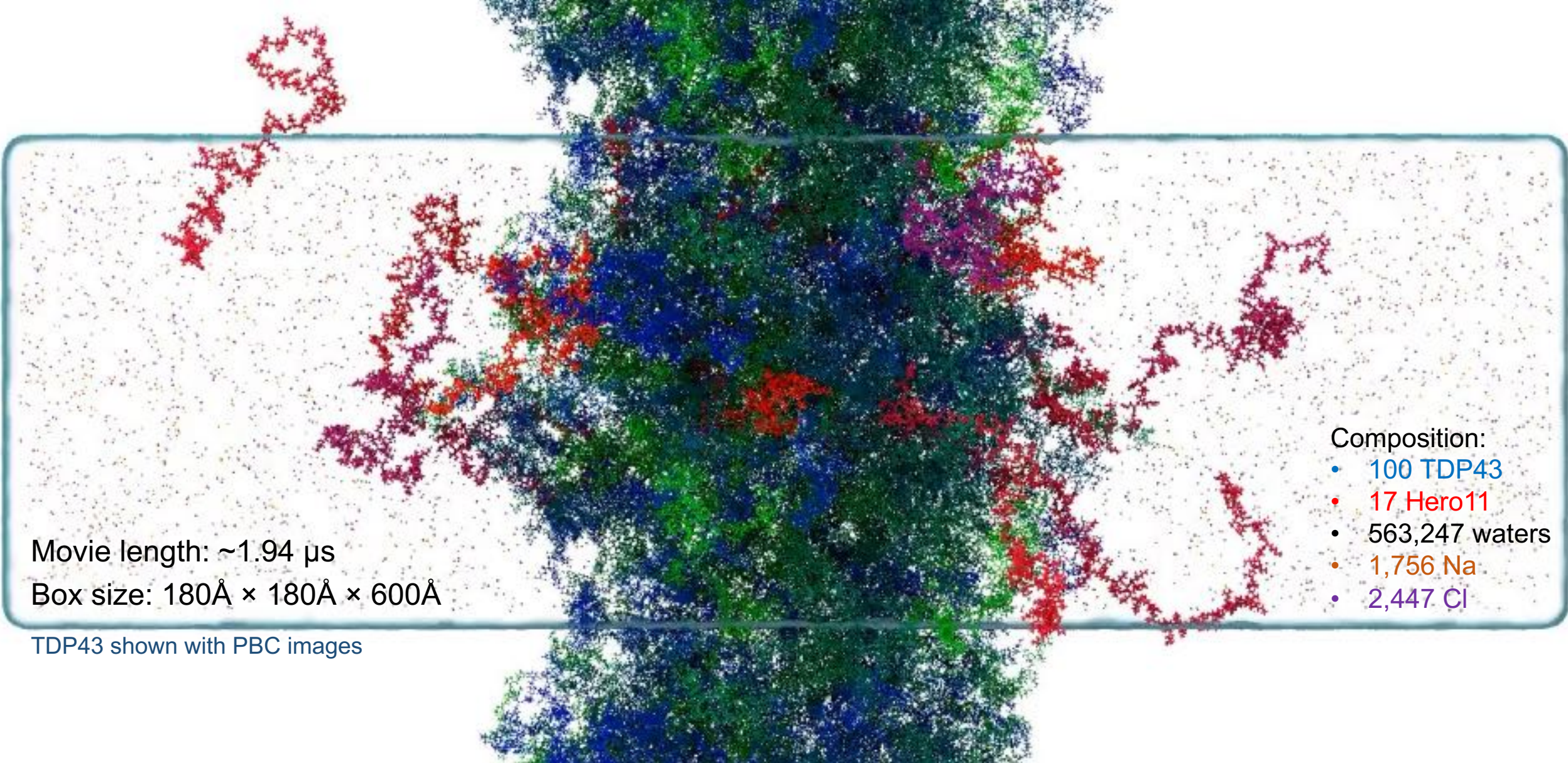
# Reconstruction of Atomistic Models from CG Simulations with GENESIS

Many tools: PULCHRA, BBQ, PD2, SCWRL...

Mori *et al.* *JCIM* (2021)



# All-atom simulations of TDP43-Hero11 Condensates



Movie length:  $\sim 1.94 \mu\text{s}$

Box size:  $180\text{\AA} \times 180\text{\AA} \times 600\text{\AA}$

TDP43 shown with PBC images

Composition:

- 100 TDP43
- 17 Hero11
- 563,247 waters
- 1,756 Na
- 2,447 Cl

# Acknowledgements

HPCIC 計算科学フォーラム

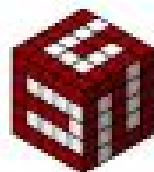
## RIKEN

- Yuji Sugita
- Jaewoon Jung
- Chigusa Kobayashi
- Diego Ugarte La Torre
- Ai Niitsu

## Shintaro Iwasaki

## U. Tokyo

- Yukihide Tomari
- Kotaro Tsuboyama
- Selena Setsu
- Andy Lam



GENESIS

Generalized-ensemble simulation system

HPCI High Performance  
Computing Infrastructure



HOKUSAI  
BIG WATERFALL

## Kyoto U.

- Shoji Takada
- Giovanni Brandani
- Azuki Mizutani

## Hokkaido U.

- Shinichi Nakagawa

## Nagahama Inst. Bio-Sci. Tech.

- Takao Yoda



非ドメイン型  
バイオポリマーの生物学