香港中文大學
The Chinese University of Hong Kong

Ab initio characterization of protein molecular dynamics with **AI^{2}BMD**

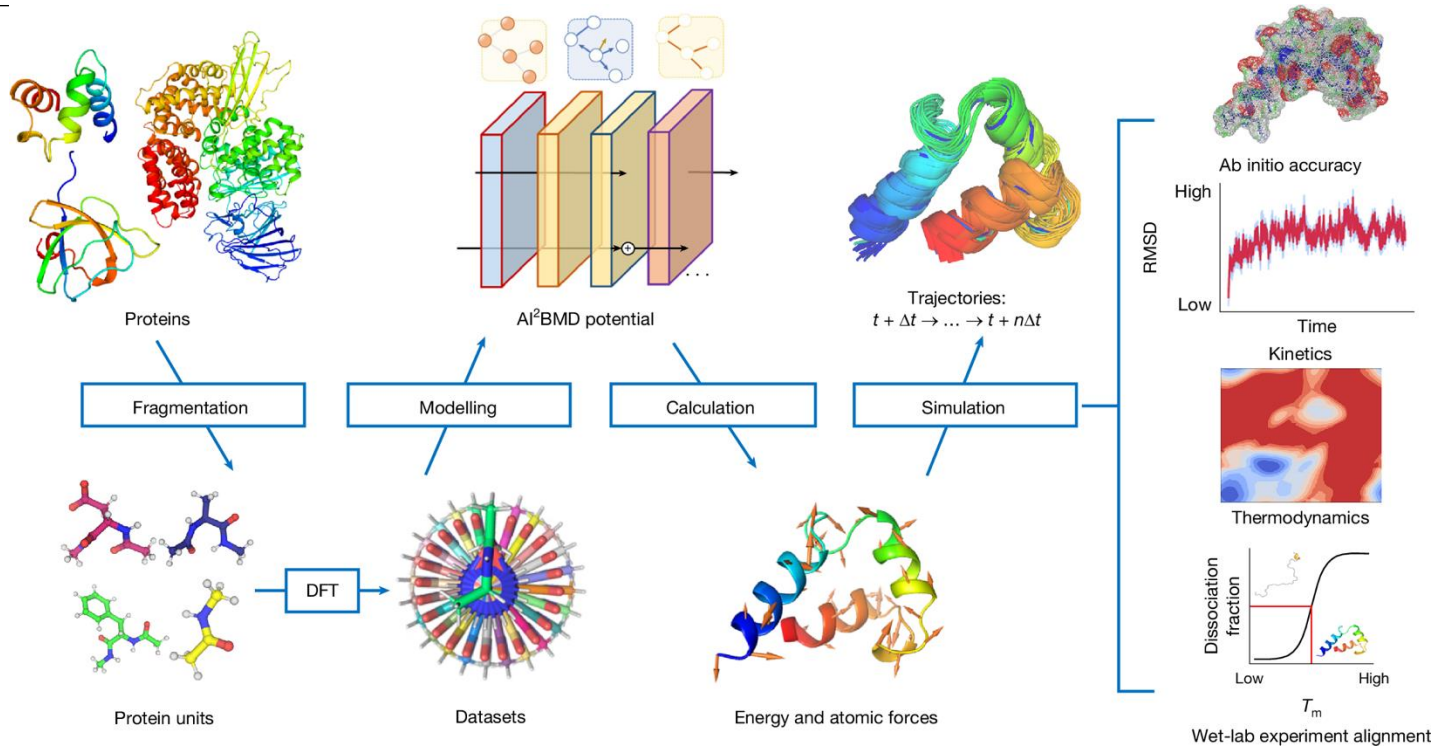
Jingyi Liu

16/01/2026

Nature 2024, 635, 1019–1027.

Outline

1. Background
2. AI²BMD workflow
3. Results and discussion
4. Conclusion and perspectives



Research Background

Research Background

- As computational simulations now achieve near-experimental accuracy, they are transitioning from auxiliary interpretative tools to a central paradigm for discovery in the life sciences.

- Molecular dynamics (MD) plays an essential role:
  - → classical MD
  - → ab inition MD (AIMD)
  - → machine learning force field (MLFFs)

Accurate but face scalability challenges

- Key challenges for biomolecular simulations:
  - → Diversity of conformational space
  - → Data scarcity due to time and cost of dataset generalization

Research Background

AI^{2}BMD **A**rtificial **i**ntelligence-based **a**b **i**nitio **b**io**m**olecular **d**ynamics system

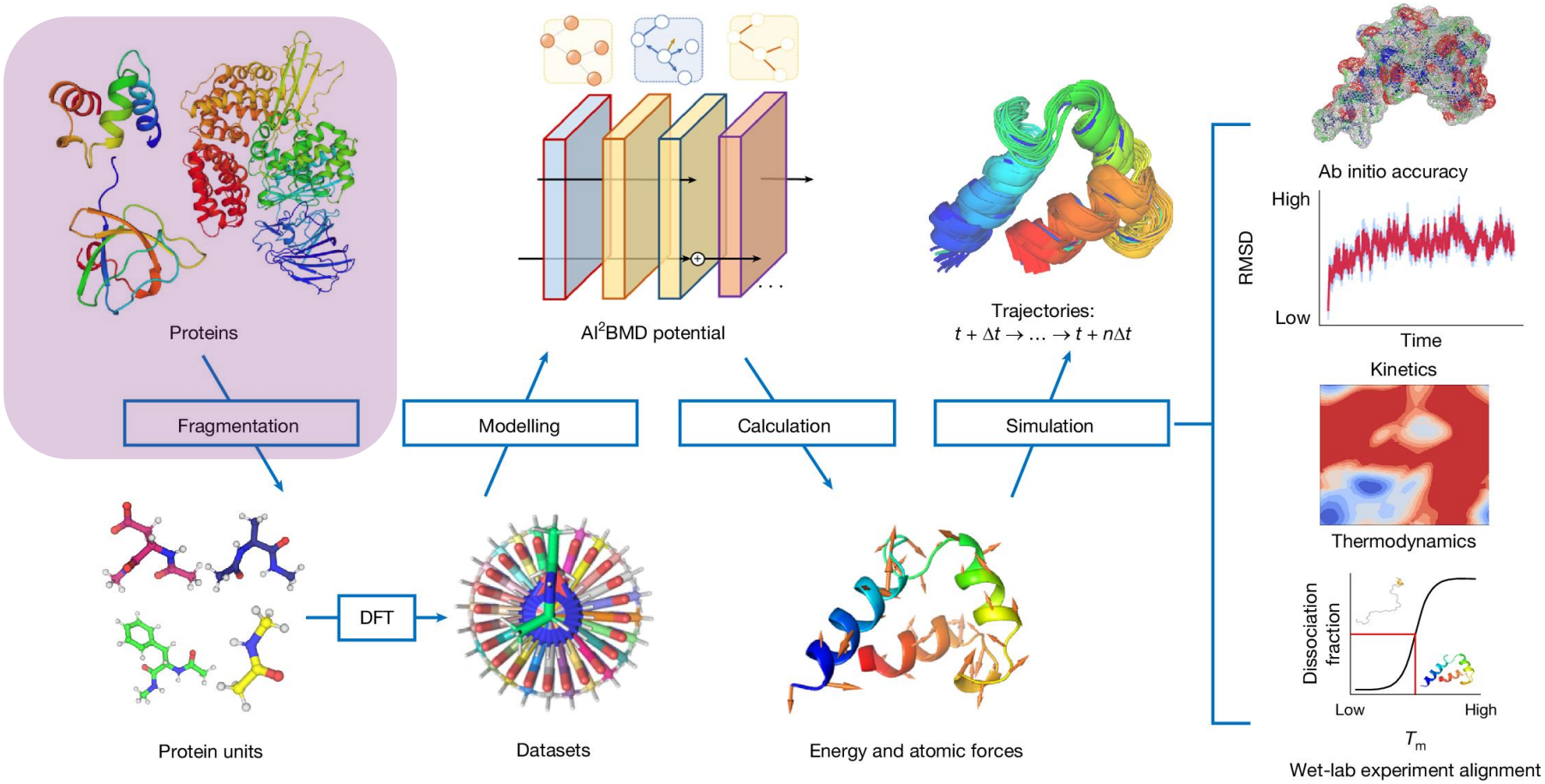
A generalizable solution for efficiently simulating a wide range of full-atom proteins with ab initio accuracy, surrounded by an explicit solvent modelled by a polarizable force field.

- A generalizable protein fragmentation approach splits proteins into overlapped protein units.
- Based on ViSNet[1], calculates the energy and atomic forces for the protein with ab initio accuracy.
- Exhibits good alignment with wet-lab experimental data, such as the melting temperature of fast-folding proteins.
- Detects different phenomena than molecular mechanics (MM)

[1] Nat. Commun. 2024, 15, 313.

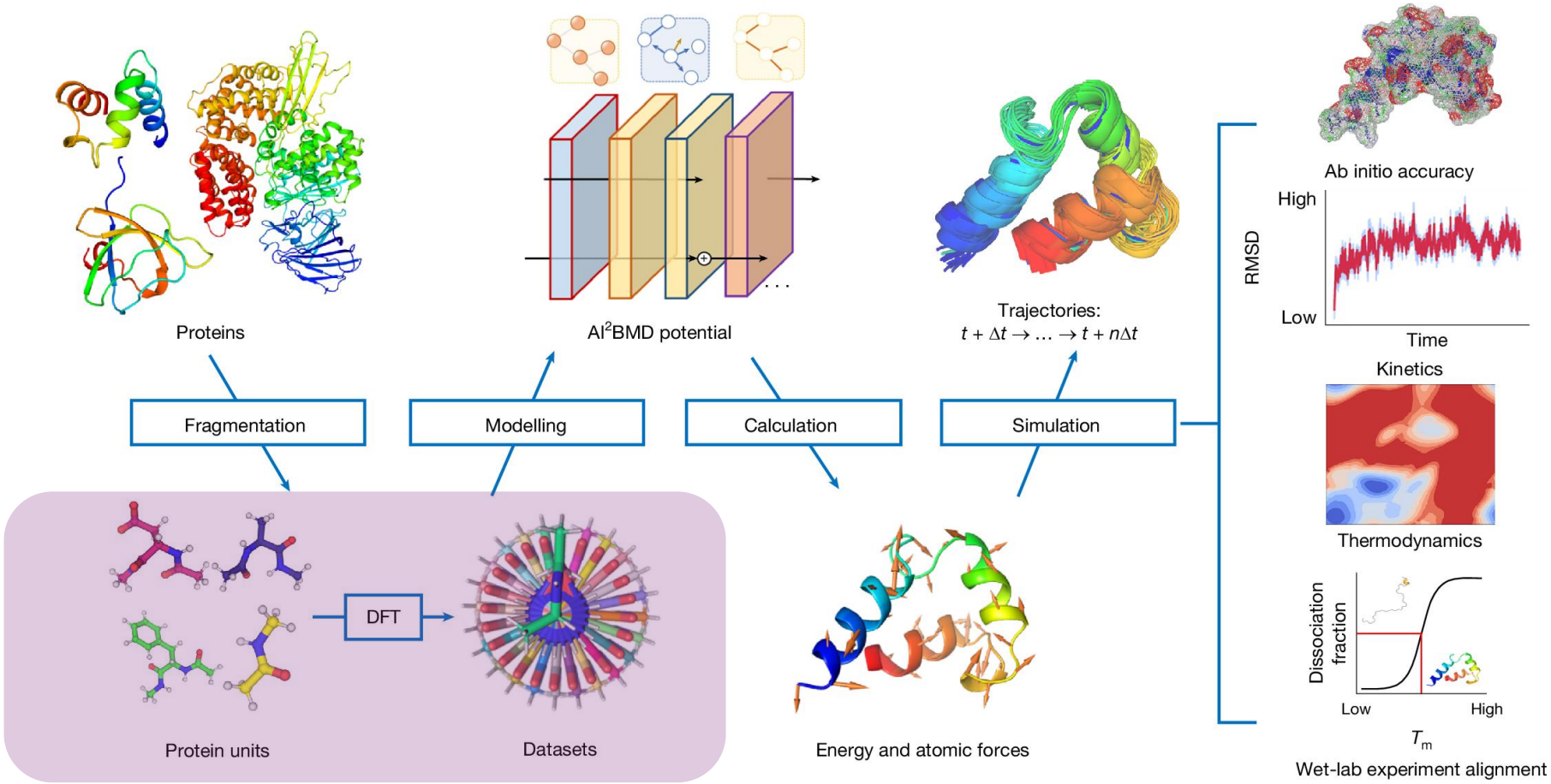
AI^{2}BMD workflow

AI²BMD workflow



1. Fragmenting proteins into smaller units, specifically dipeptides, calculate intra- and inter-unit interactions.
2. Assemble them to determine the protein energy and forces acting on the atoms

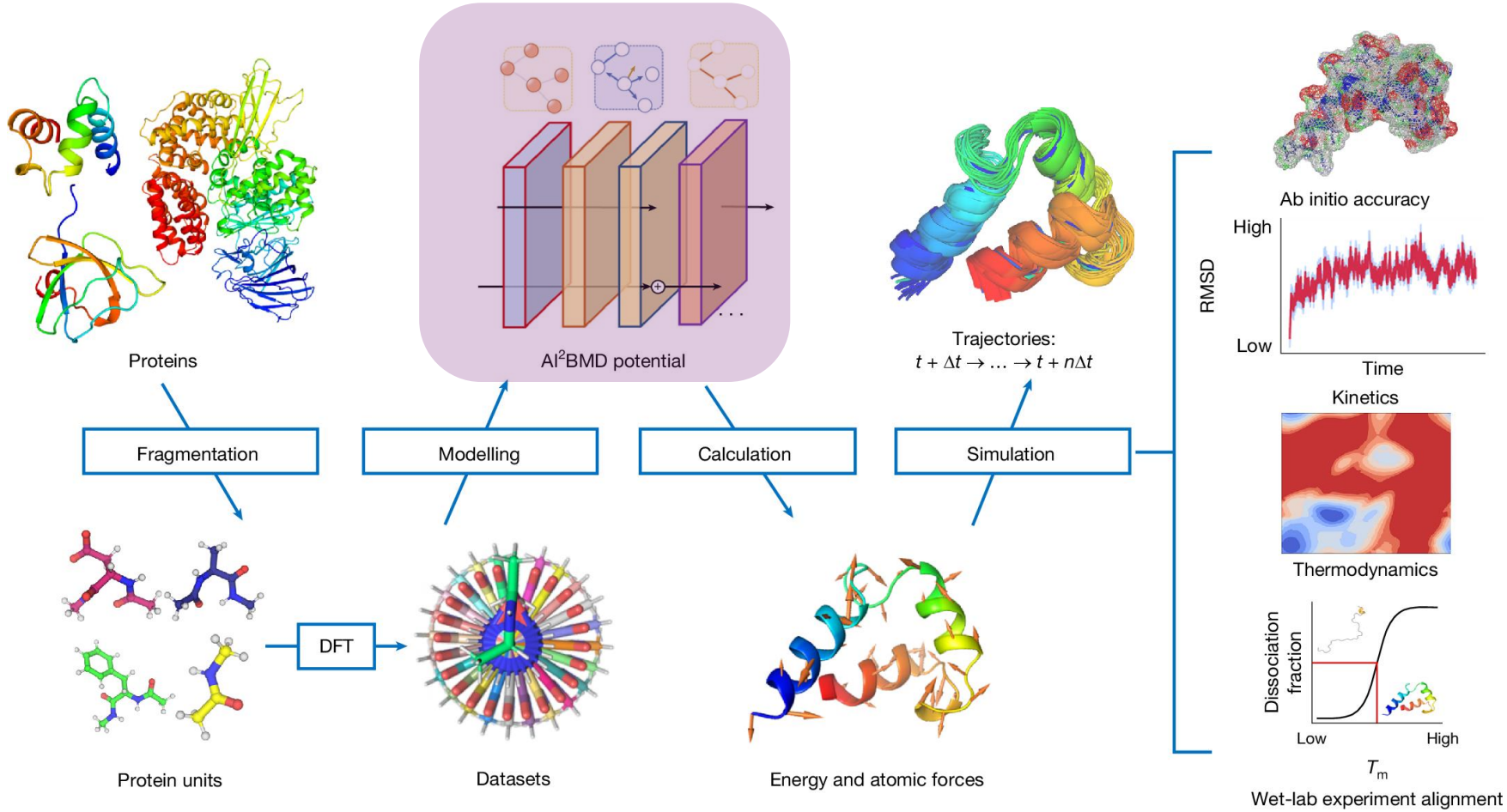
AI²BMD workflow



1. 21 protein units, with atomic number ranging from 12-36
2. Comprehensive AIMD conformation sampling by scanning main-chain dihedrals

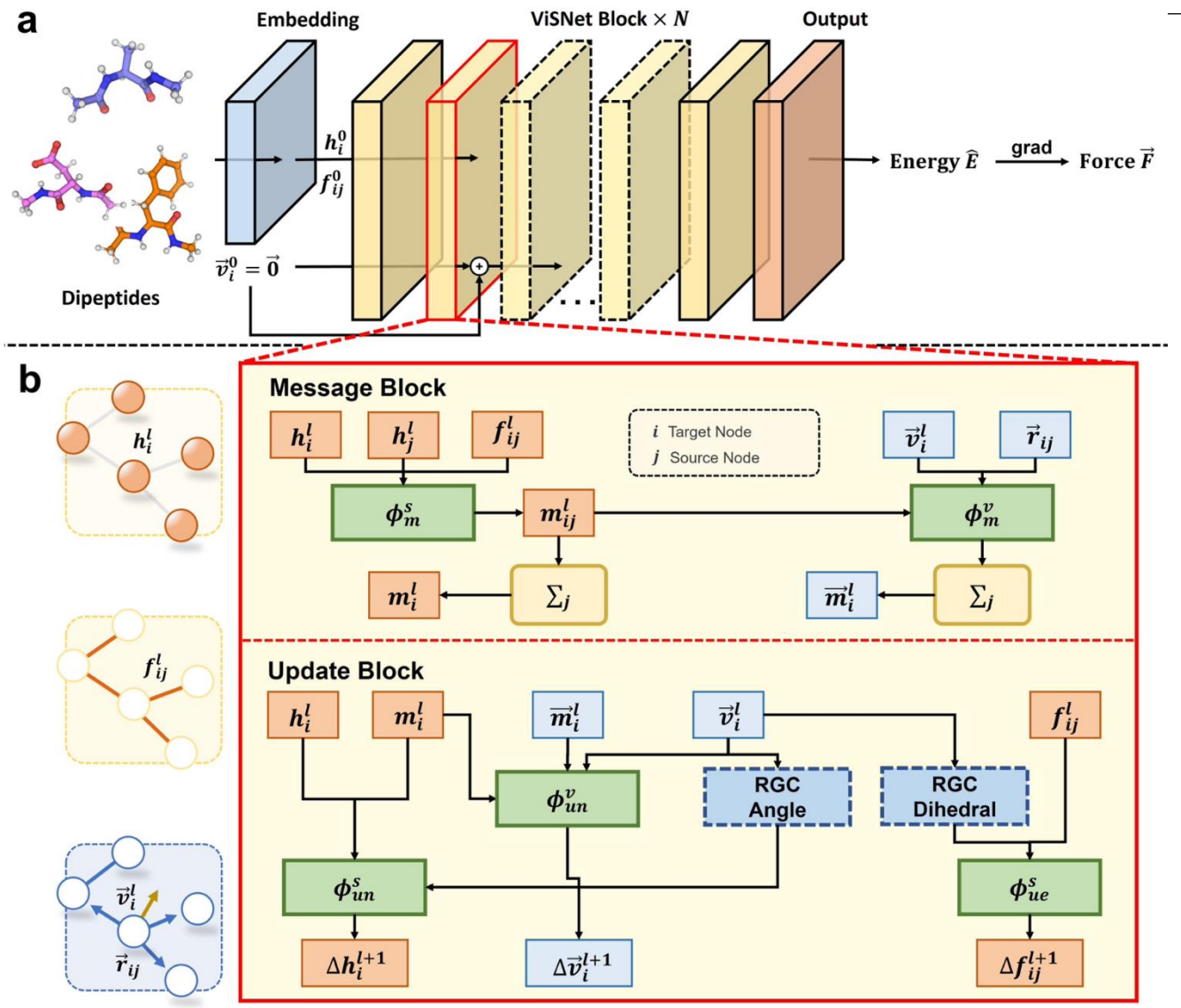
→ **20.88 million samples**

Al²BMD workflow



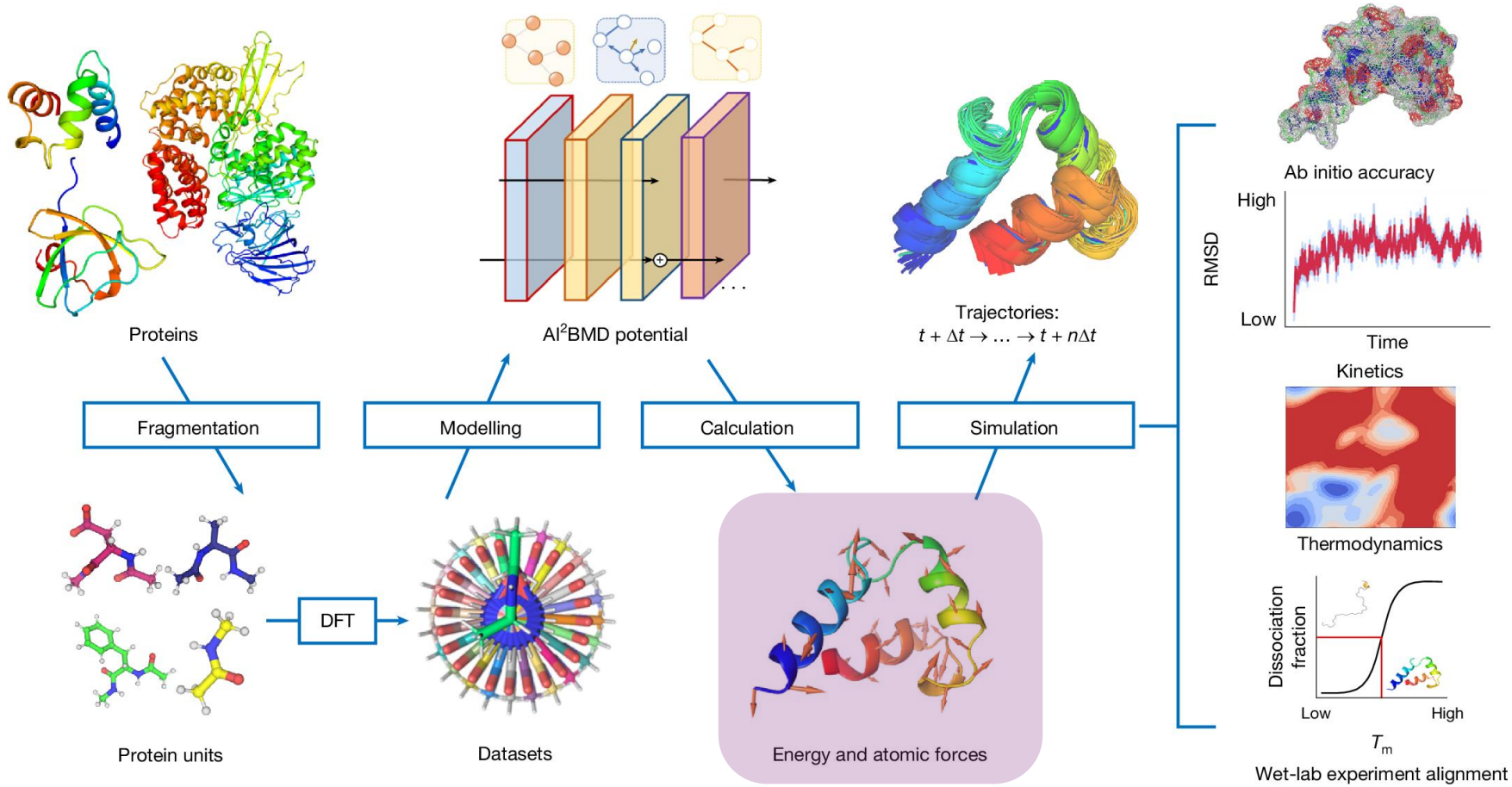
ViSNet models are trained, encoding physics-informed molecular representations and calculates four-body interactions with linear time complexity.

AI²BMD workflow



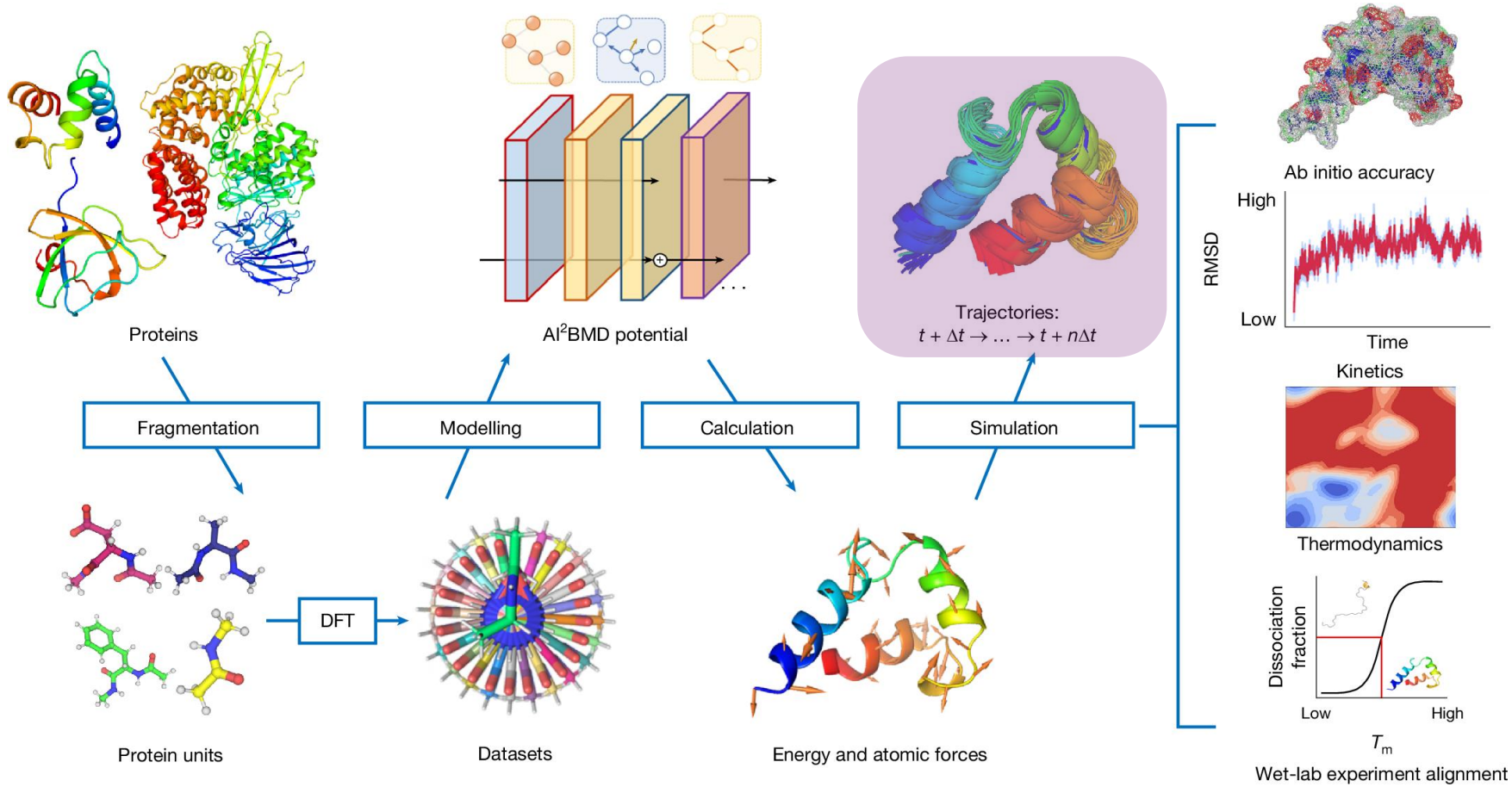
ViSNet models are trained, encoding physics-informed molecular representations and calculates four-body interactions with linear time complexity.

AI²BMD workflow



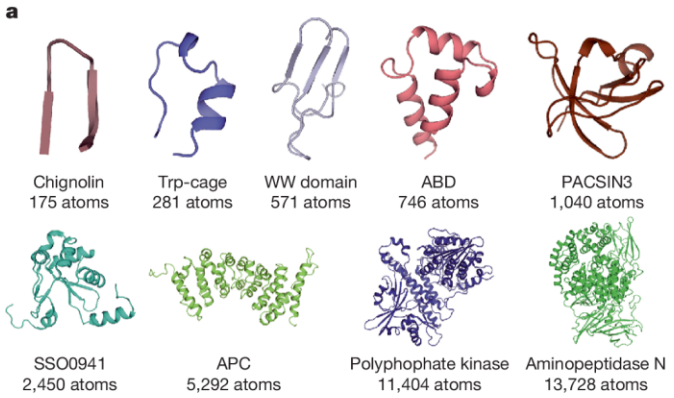
Model generates precise energy and atomic forces based on atom types and coordinates

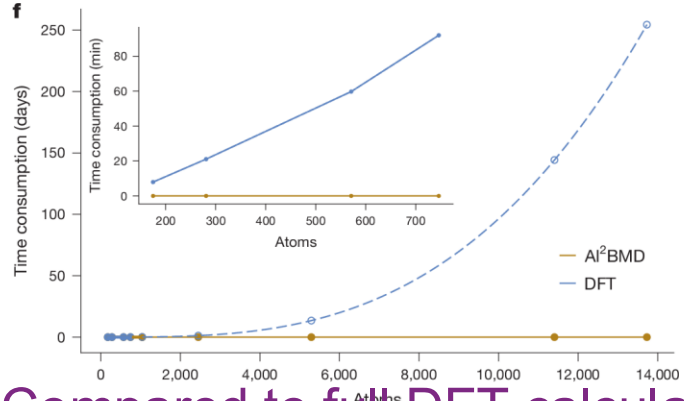
AI^{2}BMD workflow



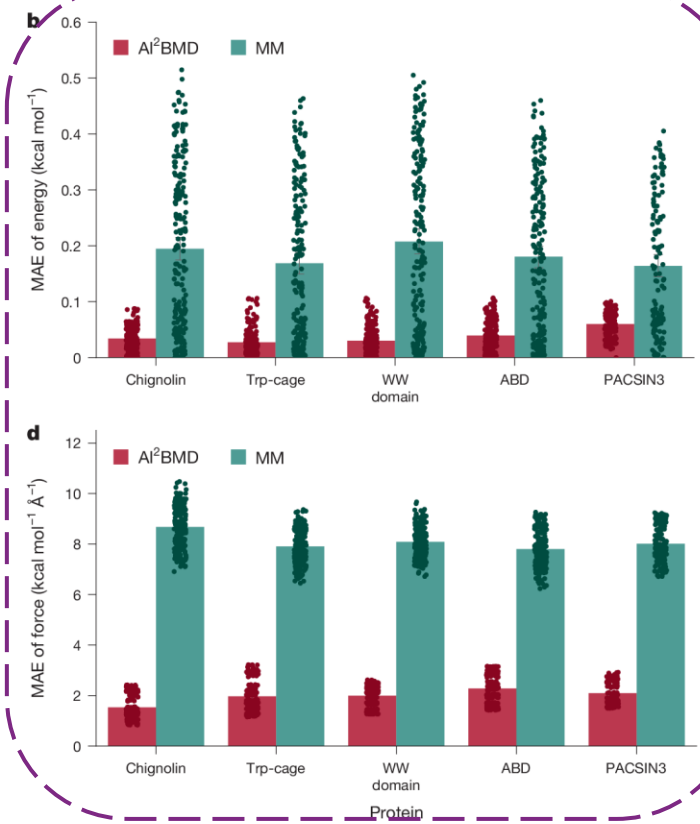
MD simulation system with polarized solvent were developed and applied for 9 proteins with number of atoms ranging from 175-13728.

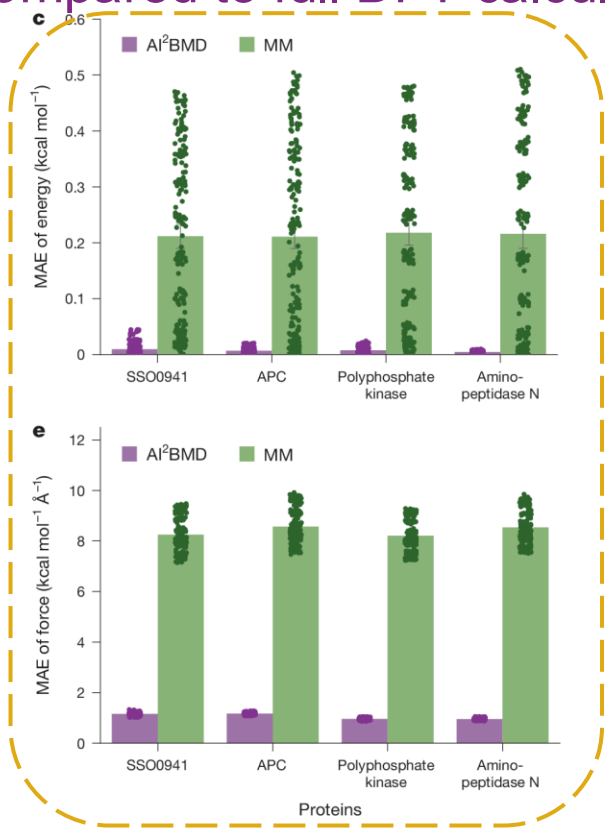
AI²BMD workflow





Compared to full DFT calculation





AI²BMD outperformed the MM force field by approximately two orders of magnitude in energy and force

(AI2BMD: 0.045 kcal mol^{-1}, 0.078 kcal mol^{-1}Å^{-1}; MM: 3.198 kcal mol^{-1}, 8.125 kcal mol^{-1}Å^{-1}).

Compared to fragment DFT calculation

AI²BMD workflow

Comparison on running time per simulation step for proteins solvated with a 10 Å water box

Protein	Atom number of protein	Atom number of system	AI²BMD (s)	DPMD (s)	Allegro (s)	Tinker Amoeba (s)	Amber FF19SB (s)
Chignolin	175	4,715	0.047	0.040	0.238	0.117	0.004
Trp-cage	281	6,067	0.052	0.055	0.322	0.136	0.005
WW domain	571	10,678	0.070	0.095	0.626	0.196	0.008
ABD	746	11,793	0.085	0.106	0.712	0.208	0.008
PACSIN 3	1,040	17,923	0.106	0.162	-	0.292	0.011
SSO0941	2,450	44,401	0.213	0.414	-	0.699	0.027
APC	5,292	54,999	0.449	0.580	-	0.938	0.033
Polyphosphate Kinase	11,404	97,657	0.966	-	-	1.487	0.058

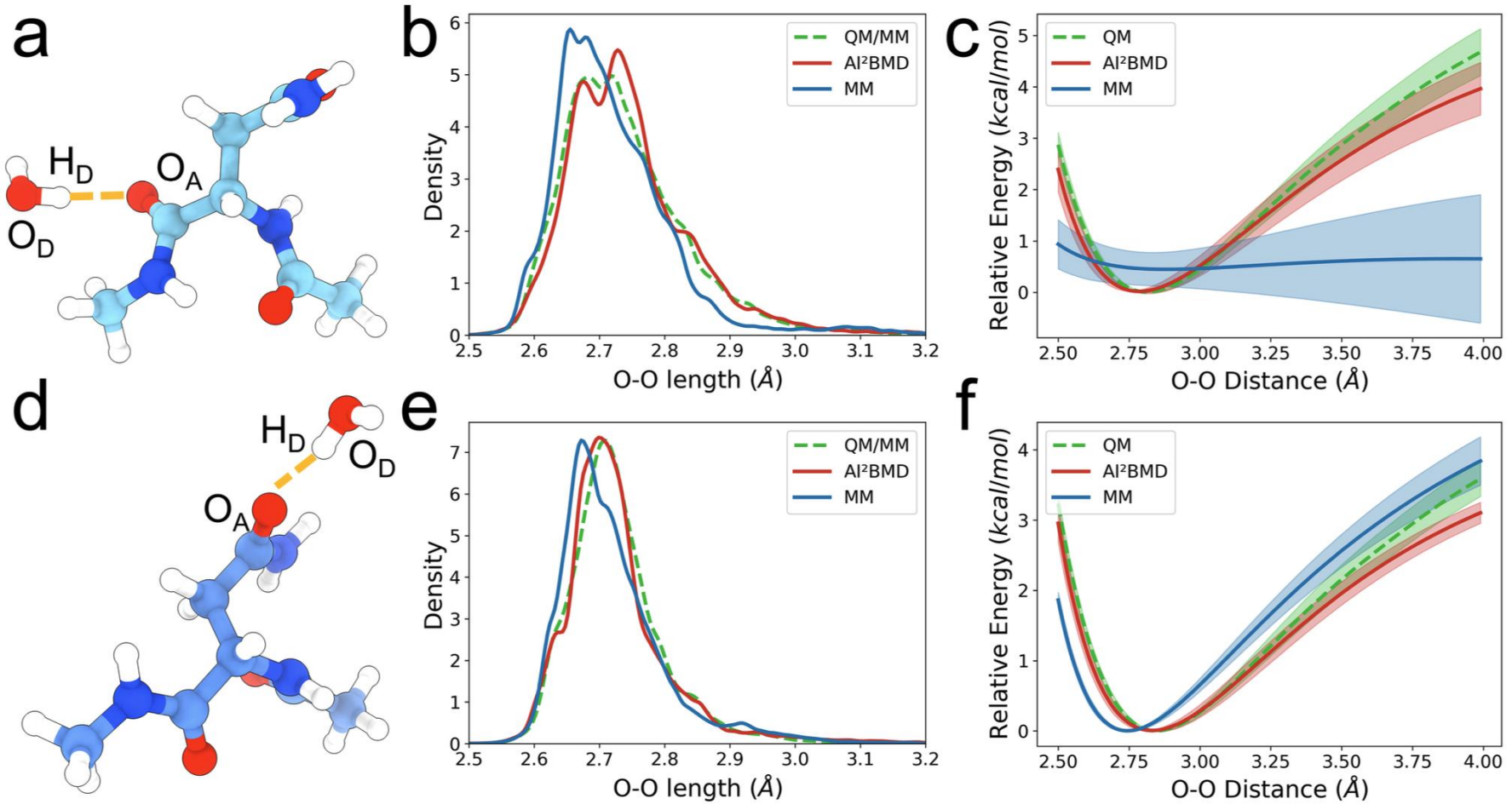
AI tools

Force-field like tools

AI²BMD outperformed other AI-driven simulation tools such as DPMD and Allegro.

AI²BMD capabilities

1. Conformational space exploration



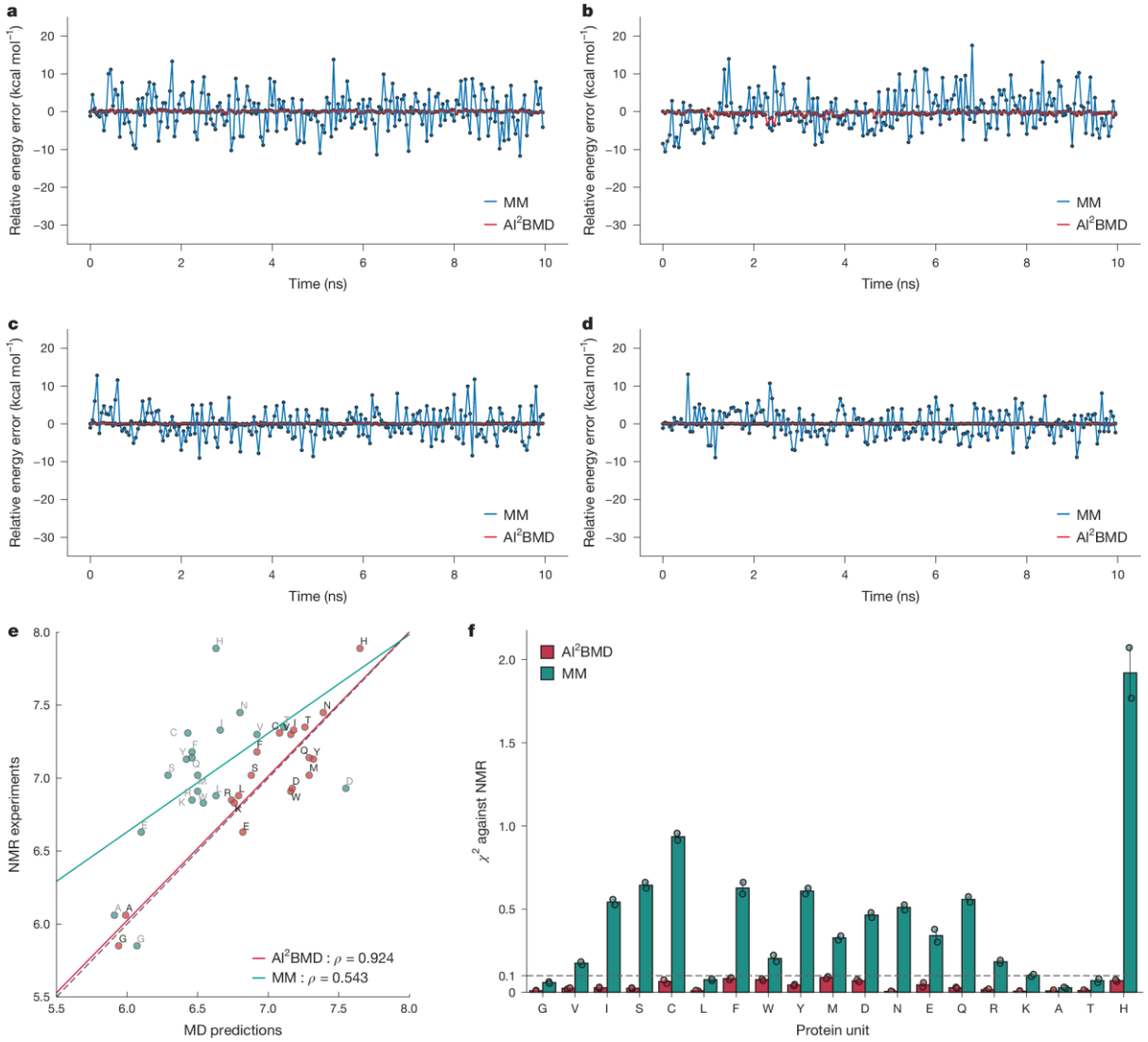
Examination of hydrogen bonding between water and the asparagine dipeptide (Ace-N-Nme) dipeptide.

AI²BMD demonstrated an energy distribution much more consistent with QM–MM than MM in bond scanning!

AI²BMD capabilities

1. Conformational space exploration

AI2BMD simulations for protein units and comparisons with NMR experiments

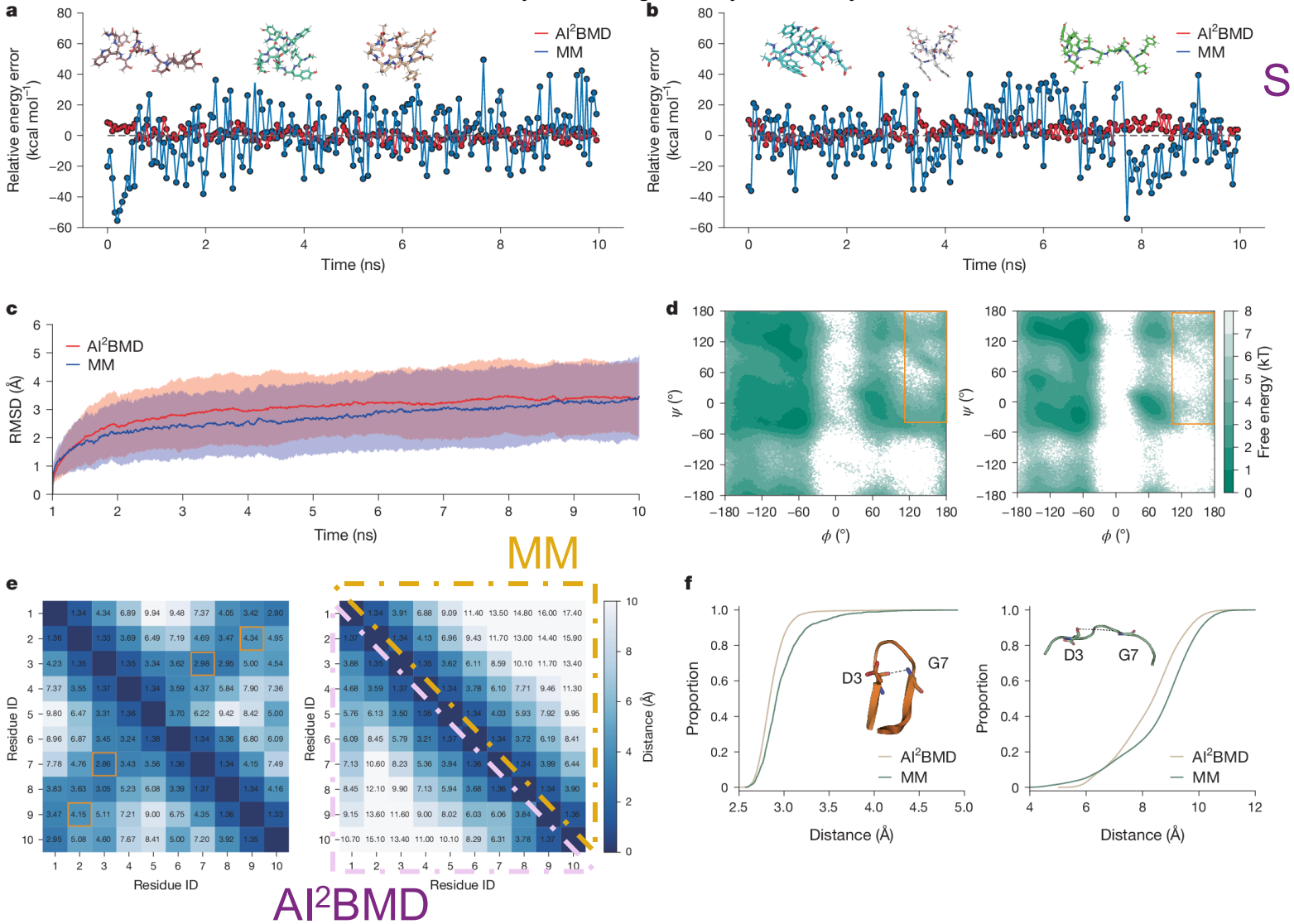


AI²BMD is much more consistent with QM–MM than MM

AI²BMD capabilities

1. Conformational space exploration

Analysis of chignolin dynamics by AI2BMD simulations



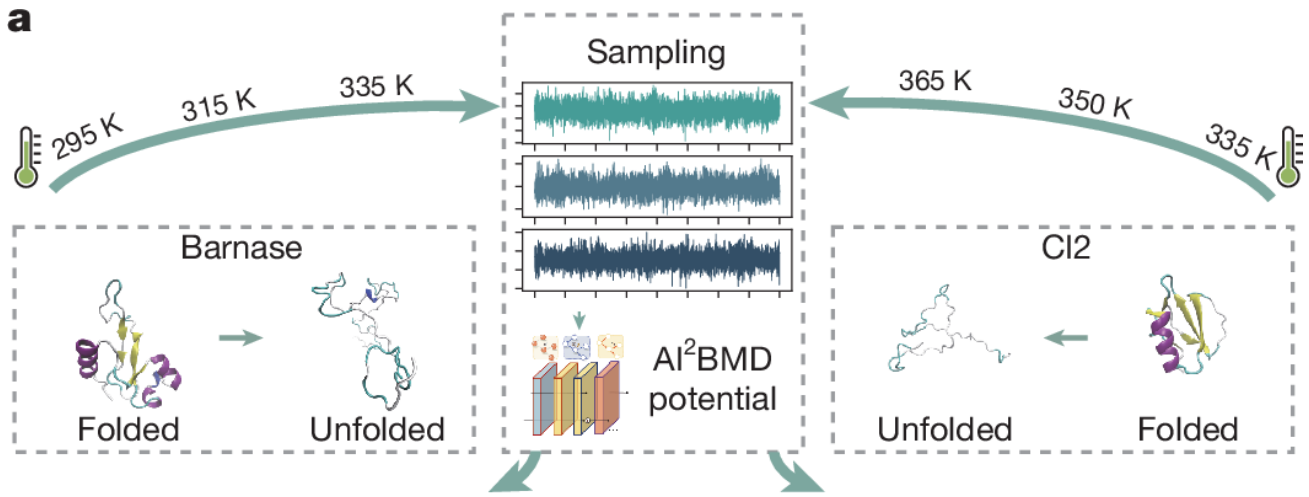
Simulations performed by AI²BMD exhibit similar structure fluctuations with MM

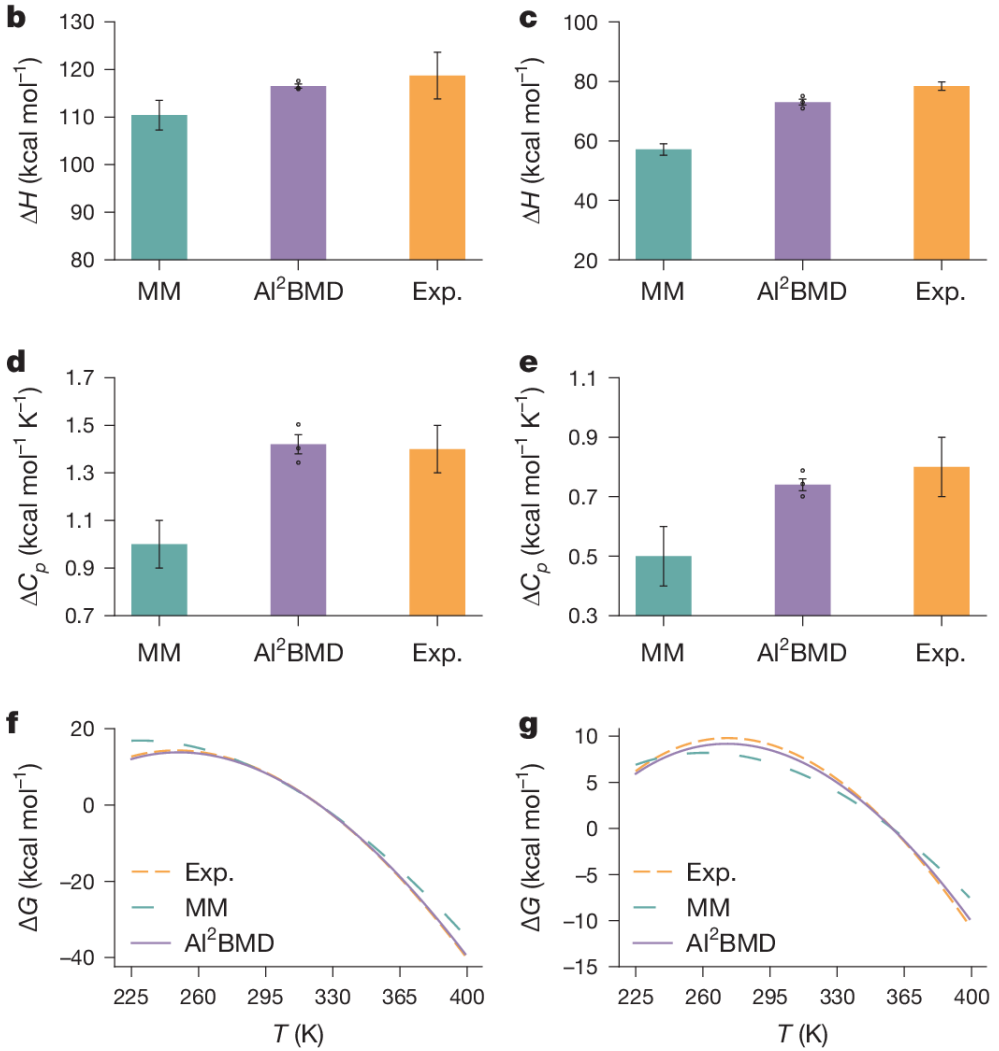
AI²BMD can detect both meaningful conformational changes and detailed interatomic interactions to study protein dynamics.

Al²BMD capabilities

2. Protein property estimation

Comparison of the change of enthalpy, heat capacity and free energy of two-state proteins, barnase and CI2.





Protein property evaluations on diverse proteins reveals Al²BMD's capability of calculating conformational ensembles accurately, leads to reasonable estimations of protein folding thermodynamic.

Conclusion and perspectives

Discussion

- AI2BMD
  - **Expands ab initio calculation from a small preset QM region to the whole full-atom protein without any prior knowledge.**
    - Eliminates the potential incompatibility of QM and MM mechanics on the boundary for proteins and accelerates QM region calculation.
    - Offer opportunities with new perspectives for complex biomolecular dynamics that QM-MM cannot deal with.
  - **Exhibits generalization ability based on fundamental assembling principles that most proteins are composed of common kinds of amino acid**, which can be expanded to other biochemical systems such as lipids, nucleotides, nanomaterials and solute-solvent interfaces.

Questions? Comments?

Thank You