Key roles of hydrodynamic interactions in protein folding

Jiaxing Yuan

Research Center for Advanced Science and Technology, The University of Tokyo

1. Introduction

Protein folding is a vital cellular process, orchestrating the creation of a complex three-dimensional structure, known as the native state, through interactions among amino acids (1). Recent strides in artificial intelligence have facilitated precise predictions of the native structure formed by specific amino acid sequences (2, 3). Despite these advancements, unraveling the non-equilibrium folding pathways to access native structures remains a formidable challenge in biophysics research (4-8). Understanding folding pathways is crucial in protein folding, where significant kinetic barriers often impede proteins from reaching their global free-energy minimum state, leading to the formation of metastable non-native intermediates (8). Proper folding is indispensable for proteins' biological functions and human health, as misfolded proteins can aggregate to form amorphous clusters or amyloid fibrils, contributing to various neurodegenerative diseases (9).

Various experimental techniques, including X-ray crystallography, fluorescence spectroscopy, nuclear magnetic resonance, circular dichroism spectroscopy, atomic force microscopy, and magnetic tweezers, have been utilized to study protein folding kinetics. These techniques have successfully unveiled the complexities of protein folding dynamics, extending the observation time to hours and days. However, achieving high spatial and temporal resolution simultaneously in experimental folding studies remains challenging. Molecular dynamics (MD) simulations offer an alternative approach, providing a detailed view of protein structure and dynamics at an extremely high spatiotemporal resolution. Nevertheless, MD simulations can be computationally expensive and are limited to timescales from microseconds to milliseconds. In contrast to atomistic simulations, coarse-grained (CG) minimal models significantly reduce computational costs, capturing essential protein folding features without delving into microscopic details. This approach has significantly contributed to a profound understanding of the fundamental mechanisms involved in protein folding.

Hydrodynamic interactions (HI) induced by solvent flow could significantly impact protein folding, dynamically coupling the motion of amino acid residues. While the roles of HI in phase ordering kinetics of soft matter systems are well-established, less attention has been given to their effects on protein folding. Experimental investigation is challenging due to the inherent incorporation of HI, making computer simulations more suitable. However, most simulations have neglected momentum conservation, such as Langevin dynamics (LD) and Brownian dynamics (BD), leading to

varied reports on the effects of HI on protein folding. To date, all CG simulations of protein folding have employed the BD method incorporating the Rotne-Prager (RP) tensor for modeling HI. However, the RP tensor cannot accurately account for the short-range many-body aspects of HI due to fluid incompressibility, critical in preventing particles from forming closely packed arrangements during protein folding, known as the "squeezing flow effect." This short-range contribution is essential for accurately modeling non-equilibrium phase ordering kinetics.

This project employs the fluid particle dynamics (FPD) simulation method [10] based on the direct computation of the Navier-Stokes equation to study the folding kinetics of a four-α-helices bundle protein. The findings suggest that HI plays a significant role in selecting fast folding pathways, preventing kinetic trapping, and accelerating folding kinetics. The directional flow expedites collapsing dynamics during protein folding, in line with prior knowledge. Importantly, incompressibility-induced squeezing flow plays a substantial role in establishing proper hydrophobic contacts, preventing the accumulation of non-native contacts and averting protein entrapment in local free-energy minima. The impact of HI is most significant for a biologically relevant quench depth, highlighting their crucial role in protein folding, beyond previous considerations based on CG simulations neglecting the short-range many-body part of HI.

2. Simulation method

Our coarse-grained (CG) model features a representative four-α-helices bundle protein positioned in a cubic three-dimensional (3D)-periodic box, incorporating many-body hydrodynamic interactions (HI) among particles using the fluid particle dynamics (FPD) method. The four-α-helices bundle protein comprises four interconnected α-helices, represented by beads of diameter σ, symbolizing three types of amino-acid residues: hydrophobic (red), hydrophilic (blue), and neutral (yellow) (Fig. 1).

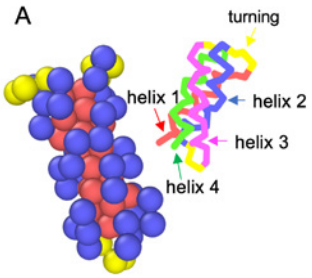

Fig1: The native conformation of four-α-helices bundle protein.

Spherical beads interact through the Lennard-Jones (LJ) potential V LJ, incorporating energy coupling ε, harmonic bond potential V_{bond}, harmonic angle potential V_{angle}, and dihedral potential $V_{dihedral}$. The folding temperature T_f is defined as $k_B T_f/\varepsilon \simeq 0.7$, where k_B is the Boltzmann constant. In Figure 1, the native structure is illustrated, obtained by gradually quenching the system to zero temperature. Folding simulations are conducted as follows: an initial configuration is prepared through equilibrium simulations under

the temperature $k_B T_{init}/\varepsilon = 1$, followed by an immediate quenching of the system to $T < T_f$.

3. Simulation Results

As a reference system, we explore the folding kinetics using BD simulations without incorporating HI (Fig. 2). We find only 3 out of 14 trajectories produce the correct folded state, characterized by $\chi \approx 0$, before $t \approx 1.2 \times 10^5\ \tau_{BD}$ (Fig. 2). We classify these successful folding pathways into two types, depending on the chronological order of folding and collapse: The collapse and folding occur almost concurrently for the type I pathway (f1 in Fig. 2), whereas in the type II pathway (f2–f3 in Fig. 2), folding is completed much later than the collapse. In the majority of instances, proteins are trapped in unfolded intermediates with $\chi \approx 0.2 - 0.4$ despite being collapsed. Among all the simulations we conducted, the quickest folding pathway has a folding time of approximately $t_f \approx 2 \times 10^4\ \tau_{BD}$. Even though the native structure of the four-α-helices bundle protein appears to be simple, the selection of indirect folding pathways with slow relaxation dynamics is quite common.

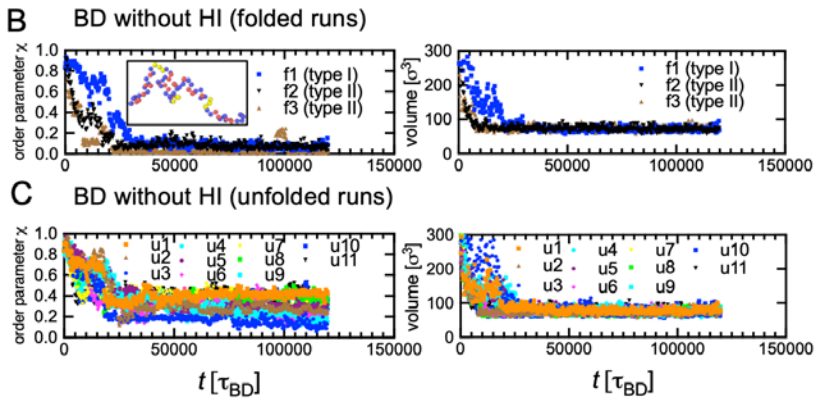

Fig2: Temporal change of order parameter χ and protein volume for folded and unfolded trajectories in BD simulations. Here the protein volume is determined by calculating the convex hull of amino-acid residues.

The situation becomes far more different when HI is included. Remarkably, we now observe that 7 of the 14 trajectories show the correct folded state. In the remaining simulations, the proteins have remained misfolded until the end of the simulation. Note, however, that due to the available computational cost, the simulation time in FPD is 7 ~ 8 times shorter compared to BD. Thus, it is possible that some of the unfolded states eventually fold after waiting the same amount of time as in BD. Thus, our results collectively show that HI helps select fast folding pathways to the native state without being kinetically trapped, speeding up the folding kinetics compared to its absence.

4. Conclusion and Outlook

In summary, we have unveiled the crucial roles of hydrodynamic interactions (HI) in the folding kinetics of a four-α-helices bundle protein employing the fluid particle dynamics (FPD) method (10). Our findings demonstrate that HI plays a pivotal role in selecting fast folding pathways, significantly accelerating folding kinetics compared to its absence.

In living cells, our findings suggest that cells could leverage HI to favor faster folding pathways, potentially reducing the likelihood of protein aggregation. Considering that many proteins and other biomolecules, such as RNA (73), have charged components influencing folding kinetics through electrostatics, it is conceivable to extend our study to charged biomolecules using the FPD method incorporating electrostatic interactions and ion dynamics. The intriguing question of how organisms have utilized HI during evolution remains, reflecting the remarkable adaptability and creativity of life that has withstood natural selection.

References

1. Lesk A (2010) Introduction to protein science: architecture, function, and genomics. (Oxford university press).
2. Jumper J, et al. (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596(7873):583-589.
3. Roney JP, Ovchinnikov S (2022) State-of-the-art estimation of protein model accuracy using AlphaFold. Phys. Rev. Lett. 129(23):238101.
4. Dobson CM (2003) Protein folding and misfolding. Nature 426(6968):884-890.
5. Thirumalai D, O'Brien EP, Morrison G, Hyeon C (2010) Theoretical perspectives on protein folding. Annu. Rev. Biophys. 39(1):159-183.
6. Freddolino PL, Harrison CB, Liu Y, Schulten K (2010) Challenges in protein-folding simulations. Nat. Phys. 6(10):751-758.
7. Gershenson A, Gierasch LM (2011) Protein folding in the cell: challenges and progress. Curr. Opin. Struct. Biol. 21(1):32-41.
8. Englander SW, Mayne L (2014) The nature of protein folding pathways. Proc. Natl. Acad. Sci.
9. Dobson CM (2002) Protein-misfolding diseases: Getting out of shape. Nature 418(6899):729-730.
10. Tanaka H, Araki T (2000) Simulation method of colloidal suspensions with hydrodynamic interactions: Fluid particle dynamics. Phys. Rev. Lett. 85(6):1338-1341.