Temperature-dependent stability and interaction changes of Rab1-DrrA complex

2137

CMML3 ICA1 mini-project 7

Abstract. Temperature critically influences host-pathogen protein complex stability and interaction dynamics. While the Rab1-DrrA complex—a key factor in *Legionella pneumophila* pneumonia pathogenesis—has been extensively studied, its behavior across temperatures remains poorly investigated. Therefore, current study presents comprehensive analyses based on molecular dynamics simulation of Rab1-DrrA complex under multiple temperature conditions. We also reveal the limitation of AlphaFold3 in comparison, highlighting its lack of temperature-dependent conformational information.

Legionella pneumophila is an intercellular human pathogen that belongs to γ-proteobacterial species, leading to a severe form of pneumonia (Wee *et al.*, 2021). To invade and manipulate host cells, through the activity of DrrA to bind to the Guanine nucleotide exchange factor (GEF)- RabGDI displacement factor (GDF) related region of Rab1, the DrrA protein of the bacteria effectively hijacks the host Rab1 protein to redirect vesicle trafficking to the Leigionella-containing vacuole, facilitating bacteria replication (Müller *et al.*, 2010).

Researchers have verified three domains of full-length (more than 600 residues) DrrA protein, and one of them are GEF/GDF domain responsible for these dual activities (Müller *et al.*, 2010). When DrrA binds to Rab1, the rearrangements in GEF/GDF domain would undergoes the most substantial transformation, with residues translocating about 28Å, while residues around 30-50 of Rab1, which was 36-41 in previous experimental complex, converting from an unstructured loop to a one-and-half turn α-helix for receiving and binding DrrA (Suh *et al.*, 2009; Figure 1a).

Temperature could be a significant factor for pathogen-host interaction to provide insight of pathogen adaptation and disease dynamic mechanism of Rab1-DrrA complex, which remains poorly studied. Therefore, in this project, DrrA fragment with GEF/GDF domain (residues 317-533) in complex with Rab1 (PDB ID: 2WWX) is explored with molecular dynamics (MD) simulation under 280K-320K (7°C-47°C), to uncover the temperature-wise influence (Suh *et al.*, 2009; Figure 1b). Temperature has not been considered in AlphaFold3, the most popular protein complex predictor (Abramson *et al.*, 2024). So in addition, we will compare structural findings with the AlphaFold3-predicted Rab1-DrrA complex, to evaluate the potential effect of the missing temperature parameter (Figure 1b).

Starting from the MD simulation step, GROMACS was used, primarily following the established protocols (Lemkul, 2024). The simulations were carried out at 280K, 300K, and 320K, with the initial input of the experimental structure 2WWX (Supplementary Figure 1a; see Methods for details). Three replicates were conducted for each temperature simulation.

Prior to detailed analysis, the quality of the simulation trajectories was confirmed (Supplementary Figure 1b-d). Comprehensive downstream analyses were then performed (Figure 1b). Root Mean Square Deviation (RMSD) of Cα atoms relative to experimental structure and first-frame production MD demonstrated structural stability, with limited fluctuations within 0.2 Å across individual chains and the whole complex (Figure 1c-d; Supplementary Figure 1e-f). The total radius of gyration (Rg) to evaluate overall compactness also remained consistent across the different temperatures, suggesting no significant folding or disintegration occurred (Figure 1e; Lobanov *et al.*, 2008).

Analysis of hydrogen bond counts within the complex (Figure 1f) and secondary structure changes (Figure 1h) revealed no correlation with temperature variations as well. Ramachandran plots indicate most residues in chains are in allowed conformational angels, but simulation in 280K and 300K identified the disallowed residues in chain Host (H), suggesting possible activating sites (Figure 1g; Supplementary Figure 2).

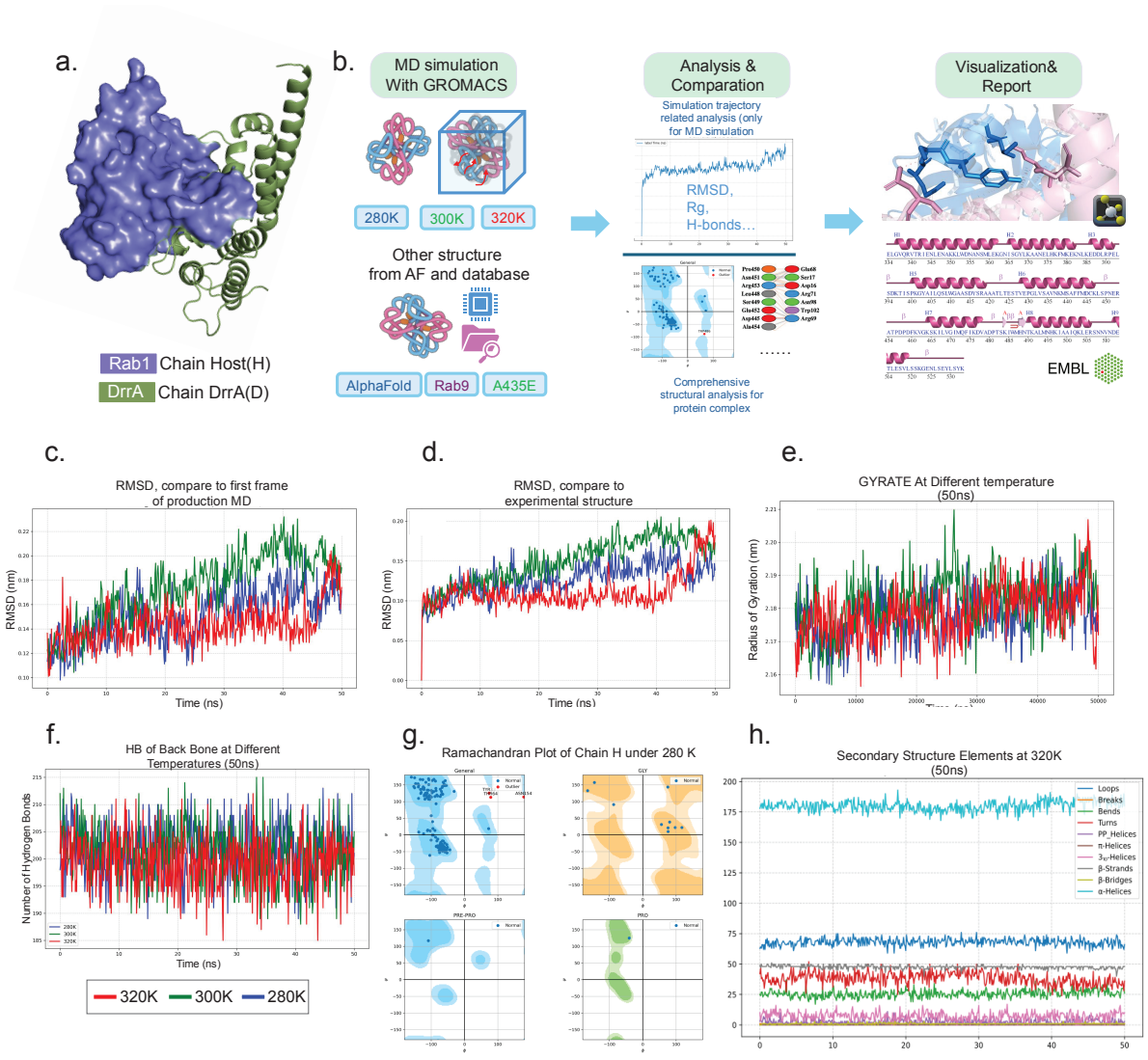


Figure 1. Comprehensive pipeline and stability assessment of simulation results (a) 3D structural visualization of Rab1-DrrA complex. (b) Pipeline of current project. Details in each step could be checked in Method section. (c) RMSD calculated by the comparison between structure in each frame and first frame of production MD. (d) RMSD calculated by the comparison between structure in each frame and experimental structure. (d) Alteration of radius of Gyration. (f) Changes of total number of hydrogen bonds of back bones in complex (g) Ramachandrane Plot generated for Chian H of 280K simulation, with several abnormal residue in disallowed region. (h) Changes secondary structure elements at 320K simulation.

Despite temperature deviations didn't significantly affect the stability, the comparison among these complexes revealed interesting findings (Figure 2a). From 280K to 300K, while simulation-to-2WWX differences of RMSD remained constant across temperatures, difference of each temperature pair was relatively high. Interface analysis indicated that higher temperatures produced more hydrogen bonds and lower Gibbs free energy, leading to stronger interactions (Figure 2b). These showed that with limited overall alteration, some specific regions, particularly the crucial binding interface, may reveal potential temperature-dependent changes.

To further identify the critical region of Rab1 for the specificity of complex interaction, Rab9, which receives another GDF PRA-1 but not DrrA, was retrieved from PDB 1WMS to compare with Rab1. 3D TM-align revealed that residues 33-44 of Rab1 was the most significantly different region for DrrA bind-

ing specificity, consistent with previous report (Suh *et al.*, 2009; Figure 2c).

This promoted a detailed examination of the significant binding interface region. Based on the detailed bonds report provided by PDBsum (Supplementary Figure2), the interaction statuses of residue 33-44 were visualized and examined (Laskowski *et al.*, 2018). Taking replicate 1 of simulation and of AlphaFold as example, interestingly, although experimental complex 2WWX has similar interaction indicators to 320K simulation, it failed to capture the active hydrogen bonds state as 320K, and missed bonds reported previously (Figure 2d; Suh *et al.*, 2009). The 320K and 300K simulations preserved expected DrrA interactions within residues 33-44, whereas 280K simulations showed significantly reduced activity with only one pair of hydrogen bonds (Figure 2d).

AlphaFold predictions slight convergence toward higher temperature simulations (Figure 2a-b; Supplementary Figure 3a-b). However, AlphaFold revealed limited performance in the inconsistent interaction patterns of replicates (Supplementary Table 1), with variable interface area and free energy, despite maintaining high chemical bond counts. Another issue was about the conflict in replicate 1, which showed hydrogen bonds comparable to 300K-320K simulations with indicator resembling 280K (Figure d). AlphaFold's limitations in generating predictions with consistent interaction properties possibly are due to its lack of temperature parameters, which are suggested to be considered in the model in the next generation.

In summary, MD simulations of Rab1-DrrA complex demonstrated thermal stability across 7-47°C (280K-320K), while the critical interface region, especially residue 33-44 of Rab1, revealed enhanced DrrA interaction in higher temperature. Lower temperature in conformational change disrupts the critical interactions between Rab1 and DrrA, particularly affecting activity of Glu-35, Trp-37, and SER-36, which are residues critical for the specificity of dual activity of DrrA for Rab1 (Figure 2e; Suh *et al.*, 2009).

This finding suggested a promising method for *Legionella pneumophila* related pneumonia treatment with peptide mimetics of Rab1 residues 33-44, which would compete effectively for the DrrA binding site at physiological body temperature (Wang *et al.*, 2022). However, other interfaces may also contribute to temperature-dependent interaction alterations. In addition, we can't explain why DrrA's binding sites at residues 33-44 of Rab1 frequently shifts, which suggested broad binding capability within the GEF/GDF domain of DrrA (Figure 2c-d). Therefore, the temperature-dependent MD simulation and analyses provided a new angle and opportunity for further investigation of Rab1-DrrA complex.

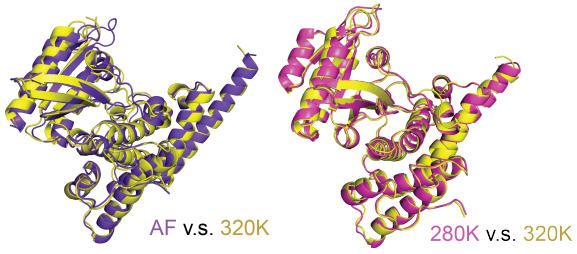
References

1. Abramson, J. *et al.* (2024) 'Accurate structure prediction of biomolecular interactions with AlphaFold 3', *Nature*, 630(8016), pp. 493–500. Available at: https://doi.org/10.1038/s41586-024-07487-w.
2. Arasaki, K., Toomre, D.K. and Roy, C.R. (2012) 'The *Legionella pneumophila* Effector DrrA Is Sufficient to Stimulate SNARE-Dependent Membrane Fusion', *Cell Host & Microbe*, 11(1), pp. 46–57. Available at: https://doi.org/10.1016/j.chom.2011.11.009.
3. Du, J. *et al.* (2021) 'Rab1-AMPylation by Legionella DrrA is allosterically activated by Rab1', *Nature Communications*, 12(1), p. 460. Available at: https://doi.org/10.1038/s41467-020-20702-2.
4. Laskowski, R.A. *et al.* (2018) 'PDBsum: Structural summaries of PDB entries', *Protein Science: A Publication of the Protein Society*, 27(1), pp. 129–134. Available at: https://doi.org/10.1002/pro.3289.
5. Lemkul, J.A. (2024) 'Introductory Tutorials for Simulating Protein Dynamics with GROMACS', *The Journal of Physical Chemistry B*, 128(39), pp. 9418–9435. Available at: https://doi.org/10.1021/acs.jpcb.4c04901.
6. Lobanov, M.Yu., Bogatyreva, N.S. and Galzitskaya, O.V. (2008) 'Radius of gyration as an indicator of protein structure compactness', *Molecular Biology*, 42(4), pp. 623–628. Available at: https://doi.org/10.1134/S0026893308040195.
7. Müller, M.P. *et al.* (2010) 'The Legionella Effector Protein DrrA AMPylates the Membrane Traffic Regulator Rab1b', *Science*, 329(5994), pp. 946–949. Available at: https://doi.org/10.1126/science.1192276.
8. Suh, H.-Y. *et al.* (2009) 'Structural insights into the dual nucleotide exchange and GDI displacement activity of SidM/DrrA', *The EMBO Journal*. Available at: https://doi.org/10.1038/emboj.2009.347.
9. Wang, L. *et al.* (2022) 'Therapeutic peptides: current applications and future directions', *Signal Transduction and Targeted Therapy*, 7(1), pp. 1–27. Available at: https://doi.org/10.1038/s41392-022-00904-4.
10. Wee, B.A. *et al.* (2021) 'Population analysis of *Legionella pneumophila* reveals a basis for resistance to complement-mediated killing', *Nature Communications*, 12(1), p. 7165. Available at: https://doi.org/10.1038/s41467-021-27478-z.

Word count Abstract: 68; Main text: 974

a.

RMSD (Å) between complexes after PyMOL alignment



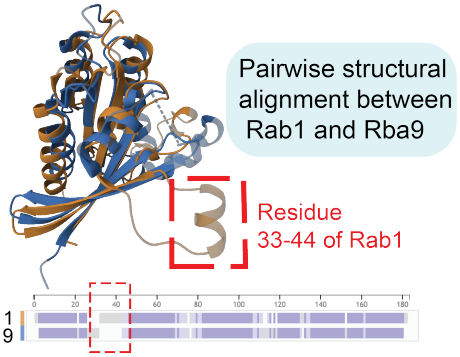
	2WWX	AF rep1	320K rep1	300K rep1	280K rep1	320K rep2
2WWX	0					
AF rep1	1.278	0				
320K rep1	0.831	1.279	0			
300K rep1	0.908	1.419	1.132	0		
280K rep1	0.934	1.681	1.195	1.02	0	
320K rep2	0.913	1.413	0.862	1.261	1.394	0

b.

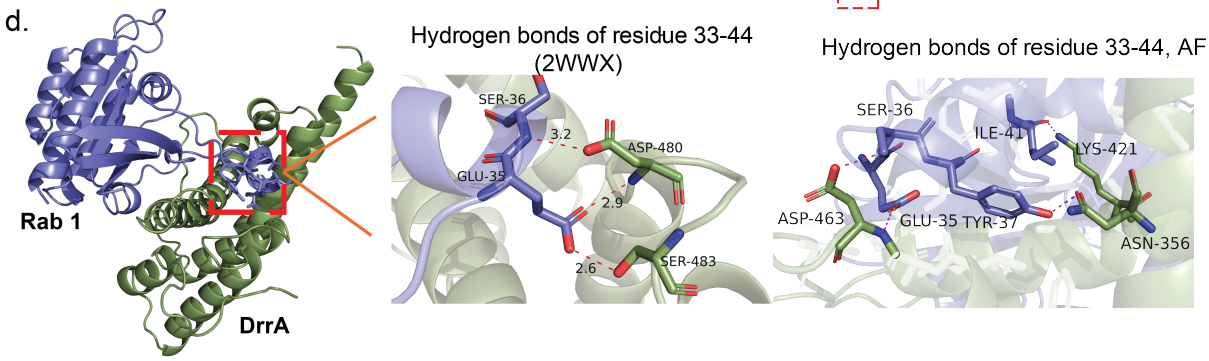
Statistic summary of interface interaction of DrrA-Rab1 *Average of replicates

Object	Interface area, Å^2	ΔG kcal/mol	No. of Hydrogen Bonds	No. of Salt Bridges	No. of non-bounded contacts
2WWX	1693.6	-13.4	24	3	247
AF*	1719.7	-11.6	22.3	3	216.7
320K sim*	1689.5	-13	24.7	2.3	217.3
300K sim*	1595.1	-10.4	21	2.7	175.7
280K sim*	1608.1	-8.8	20.7	4	186.3

c.



d.



e.

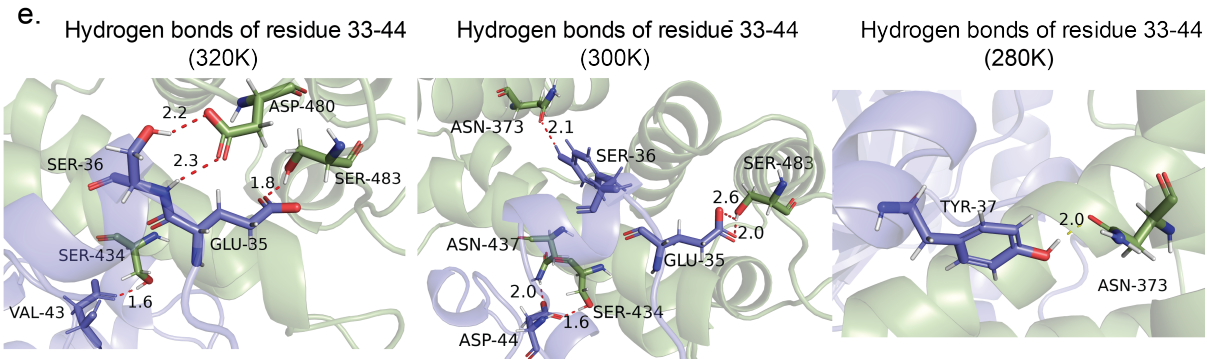


Figure 2. Structural comparison and interface analysis, with specific visualization of residue 33-44 region of Rab1 (a) RMSD calculated by comparing structures among experimental structure, AlphaFold predication, and MD simulation results. The two subplots reveal the 3D structural differences. (b) Summery of interface interaction indicators of Rab1-DrrA. Some indicators were calculated by averaging. Interface area and free energy were analysed by PISA, while number of chemical bonds were calculated by PDBsum (see Method). (c) Identifying the different region between Rab1 and Rab9. (d,e) The detailed visualization focused on the interface of residue 33-44 of Rab1, interacting with DrrA.

Supporting Materials: Temperature-dependent stability and interaction changes of Rab1-DrrA complex

Method

Molecular dynamic simulation framework

This study employed GROMACS version 2024.3 for all molecular dynamics (MD) simulations in a Linux sever of Prof. Hugo Samano's lab. The simulation framework followed GROMACS official tutorials, published methodologies (Lemkul, 2024), and example implementations from CMML3 Workshop with necessary modifications (Supplementary figure 1a)

System preparation began with protein topology definition. We utilized the CHARMM36m force field (Huang *et al.*, 2017) with default water model. Next, the system defined the periodic cubic cell with solvent addition. The system also need to add proper amount of salt into the system for neutralization (Lemkul, 2024). After these preparation, molecular mechanics calculation energy minimization (EM) would be performed. Having reached a low potential energy, with multiple temperature setting, canonical (NVT) equilibration and isothermal–isobaric (NPT) equilibration would be continued. Finally, the unrestrained production MD simulation lasting 50 ns would be conducted.

GROMACS processed the reimaging of trajectory, root-mean-square deviation (RMSD) of backbone, Radius of gyration (Rg), and hydrogen bond dynamics, changes of secondary structures (DSSP) (Carugo and Pongor, 2001; Lobanov, Bogatyreva and Galzitskaya, 2008; Gorelov et al., 2024). A new python tool has been developed and is available in GitHub, which can read XVG file of GROMACS and output images or PDF (see Code availability).

Predicted structure from AlphaFold3

AlphaFold, particularly its latest version AlphaFold3, has protein structure prediction with remarkable accuracy. AlphaFold3 sever of DeepMind was used (https://alphafoldserver.com/).

Structural and interface analysis

RCSB PDB provided online analysis for TM-align score calculation for structural alignment between protein pairs and structural difference identification (Zhang and Skolnick, 2005), responsible for results of Figure 2a.

For macromolecular complex stability assessment, we employed PDBePISA (PISA) to determine interface areas and binding Gibbs free energies (ΔG) (Krissinel and Henrick, 2007). For detailed interactions across interfaces, PDBsum was used, providing status of salt bridges, hydrogen bonds, and non-bonded contacts (Laskowski *et al.*, 2018).

Current plotting method for Ramachandran plots, including previous PyRAMA and online service of PROCHECK, could not generate separate report for each chain in a complex (Laskowski *et al.*, 1993). Therefore, a customized tool forked from PyRAMA has been developed, including new features for automatically multiple chain detection and visualization is necessary for complex PDB analysis, and it will be available for researcher via PyPi soon.

PyMOL 3D structural visalization

All molecular visualizations and structural representations were created using PyMOL version 2.6.2 in MacOS. Its capabilities allow for automated rendering of trajectory snapshots, highlighting specific interactions, and creating publication-quality images. The hydrogen bonds took results from PDBsum as references. The function of 'align' two complex structures and then calculated RMSD were utilized in results of Figure 2a. Meanwhile, PyMOL was used for assembling trajectory video for each simulation.

Replications

Three replicates have been conducted in MD simulation and AlphaFold generation, to reduce the influence of stochastic process. Unless otherwise specified, the visualization in figures used result from replicate 1.

Reflection

Mini project 7 provided an excellent opportunity to practically apply the CMML3 protein modeling curriculums, covering PDB analysis, GROMACS MD simulations, PyMOL visualization, AlphaFold, etc. Solid foundation in Python and shell coding abilities was essential for this task as well. Critical literature reviews for this project allow me to expand the

boundary of my knowledge. Still, the real challenges lay in every specific problem of my assigned complex, demanding independent exploration and leading to a customized pipeline beyond just following standard protocols.

In addition, strong initiative in developing new bioinformatic tools is valued in this project. Facing limitations of previous methods, I created a XVG visualizer, and enhanced PyRMMA for multi-chain complexes, both planned for future open-source release. Beyond the guidance, I explored the difference among Rab families, which lead to identification of significant DrrA-specific binding region of Rab1 (Figure 2c).

In the future, the techniques skills could be applied more than protein interaction. Molecular modeling, simulation, and tool development are also essential for RNA-protein interaction exploration, drug discovery, and enzyme engineering, solidifying my computational biology skillset and prepared me for diverse future challenges.

Data availability

Published experimental structure file of Rab1 complex could be obtained from `https://www.ebi.ac.uk/pdbe/entry/pdb/2wwx/index`, and the related article is available at `https://www.embopress.org/doi/full/10.1038/emboj.2009.347`. The experimental structure of Rab9 can be found in `https://www.rcsb.org/structure/1WMS`. The complexes generated from AlphaFold3, structure after GROMACS simulation (`https://alphafoldserver.com/`) and all relative analysis results have been uploaded to GitHub repository `https://github.com/AnonymityICAuser/CMML3_additional_data`.

Code availability

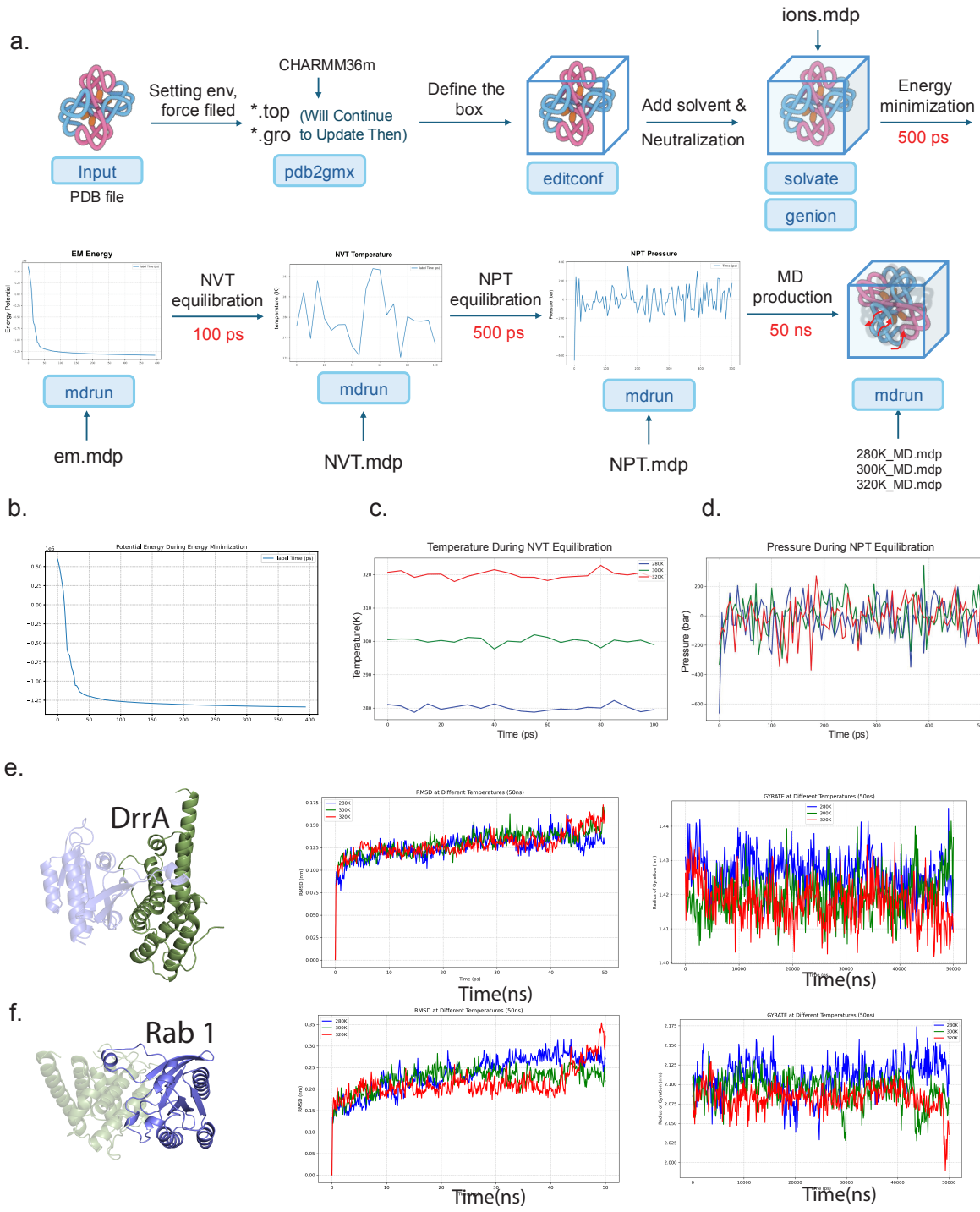
PISA online service is available at `https://www.ebi.ac.uk/pdbe/pisa/`, and PDBsum is available at `https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/`. TM-align calculation could be accessed via RSCB PDB `https://www.rcsb.org/alignment`.

All critical scripts used to run MD simulation and local analysis are available in the GitHub repository at: `https://github.com/AnonymityICAuser/CMML3_MD_simulation`. The detailed parameter settings and technique details are available in the related GitHub Document. Public PyRAMA version is in a GitHub repository and could be obtained from PyPi (`https://github.com/gerdos/PyRAMA`), but new modified and enhanced version is recommended, which could be accessed in GitHub repository `https://github.com/AnonymityICAuser/PyRAMA_enhanced`

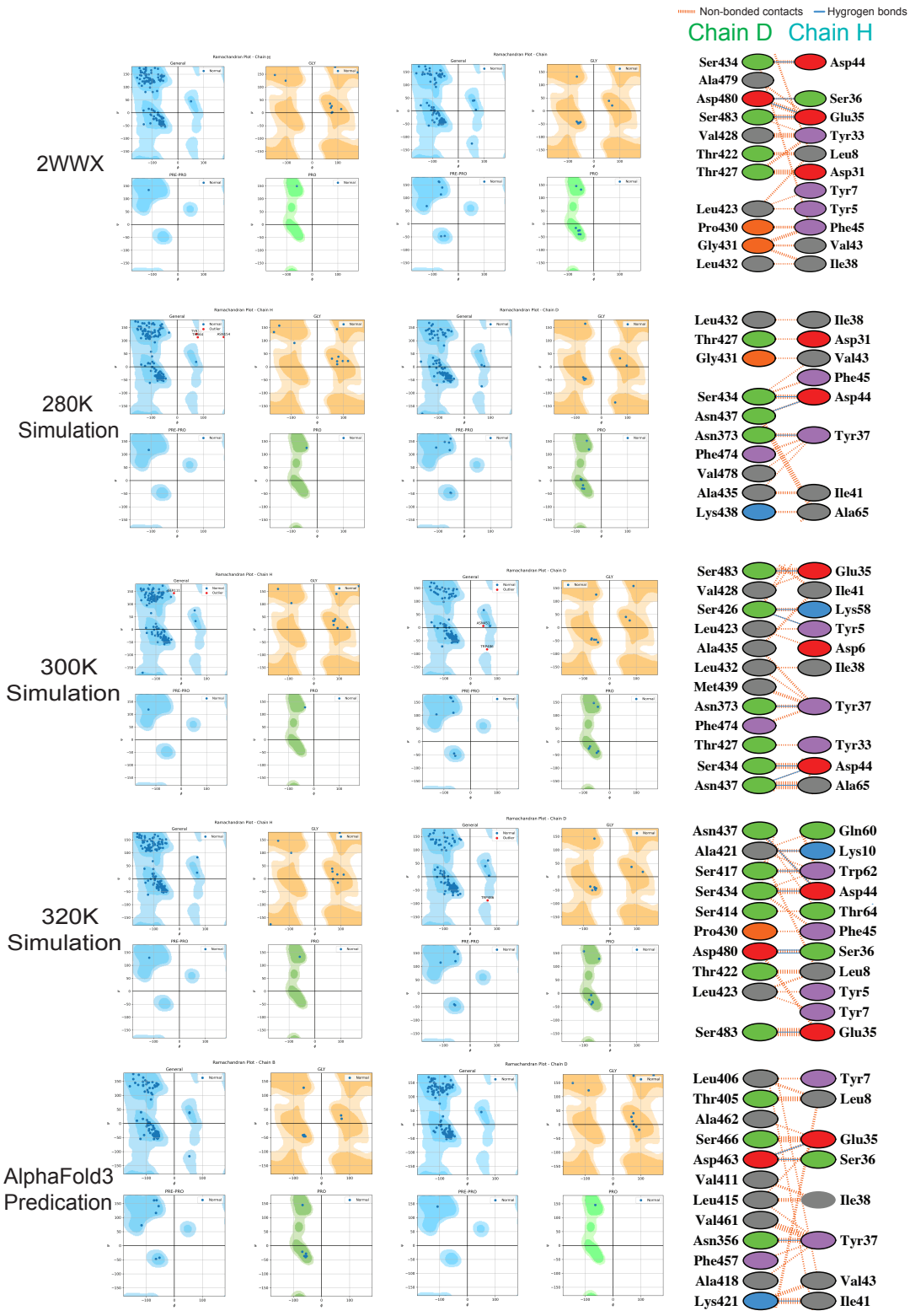
References

1. Abraham, M.J. *et al.* (2015) 'GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputer', *SoftwareX*, 1–2, pp. 19–25. Available at: `https://doi.org/10.1016/j.softx.2015.06.001`.
2. Abramson, J. *et al.* (2024) 'Accurate structure prediction of biomolecular interactions with AlphaFold 3', *Nature*, 630(8016), pp. 493–500. Available at: `https://doi.org/10.1038/s41586-024-07487-w`.
3. Carugo, O. and Pongor, S. (2001) 'A normalized root-mean-spuare distance for comparing protein three-dimensional structures', *Protein Science : A Publication of the Protein Society*, 10(7), pp. 1470–1473.
4. Gorelov, S. *et al.* (2024) 'DSSP in GROMACS: Tool for Defining Secondary Structures of Proteins in Trajectories', *Journal of Chemical Information and Modeling*, 64(9), pp. 3593–3598. Available at: `https://doi.org/10.1021/acs.jcim.3c01344`.
5. Krissinel, E. and Henrick, K. (2007) 'Inference of Macromolecular Assemblies from Crystalline State', *Journal of Molecular Biology*, 372(3), pp. 774–797. Available at: `https://doi.org/10.1016/j.jmb.2007.05.022`.
6. Laskowski, R.A. *et al.* (1993) 'PROCHECK: a program to check the stereochemical quality of protein structures', *Journal of Applied Crystallography*, 26(2), pp. 283–291. Available at: `https://doi.org/10.1107/S0021889892009944`.
7. Laskowski, R.A. *et al.* (2018) 'PDBsum: Structural summaries of PDB entries', *Protein Science : A Publication of the Protein Society*, 27(1), pp. 129–134. Available at: `https://doi.org/10.1002/pro.3289`.
8. Lemkul, J.A. (2024) 'Introductory Tutorials for Simulating Protein Dynamics with GROMACS', *The Journal of Physical Chemistry B*, 128(39), pp. 9418–9435. Available at: `https://doi.org/10.1021/acs.jpcb.4c04901`.
9. Lobanov, M.Yu., Bogatyreva, N.S. and Galzitskaya, O.V. (2008) 'Radius of gyration as an indicator of protein structure compactness', *Molecular Biology*, 42(4), pp. 623–628. Available at: `https://doi.org/10.1134/S0026893308040195`.
10. Zhang, Y. and Skolnick, J. (2005) 'TM-align: a protein structure alignment algorithm based on the TM-score', *Nucleic Acids Research*, 33(7), pp. 2302–2309. Available at: `https://doi.org/10.1093/nar/gki524`.

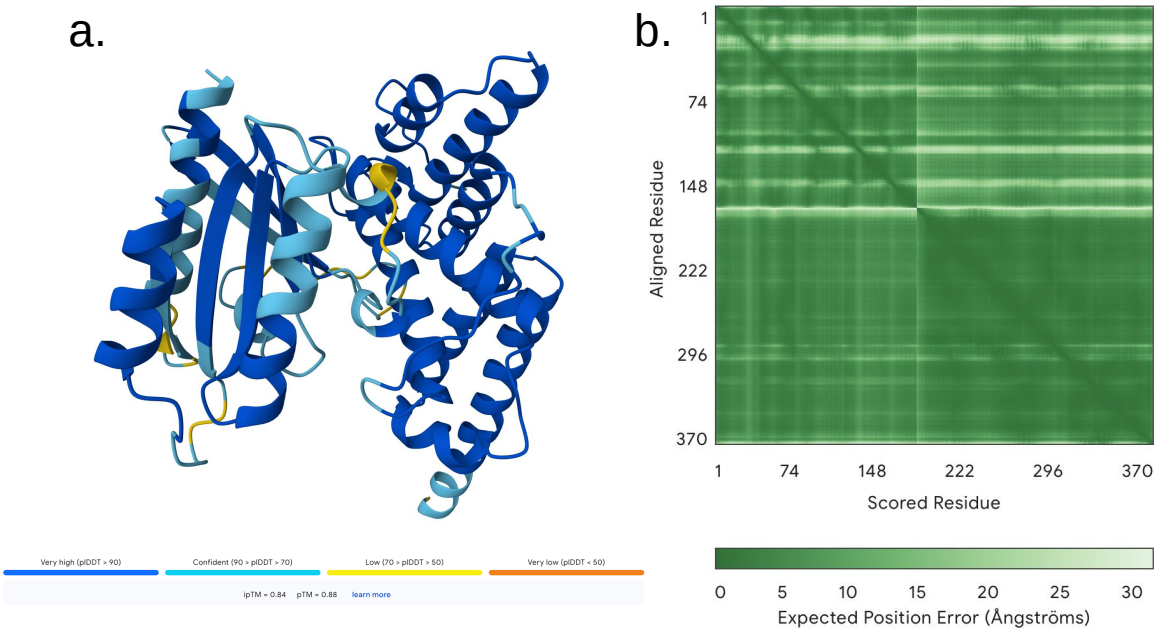
Word count Method: 494; Reflection: 192



Supplementary Figure 1. The detailed MD simulation pipeline and extended evidence for demonstrating the availability of simulation (in replicate 1). (a) GROMACS usage workflow with practical parameter setting and runtime. (b) Evaluation of potential energy during energy minimization. It showed that the energy had reached the relative minimized value. (c) Changes in temperature during NVT equilibration. (d) Changes in pressure during NPT equilibration, with the targeted pressure in 0 bar. (e) RMSD and Rg calculation of individual DrrA. (f) RMSD and Rg calculation of individual Rab1. Two individual protein both revealed stability in all temperature condition.



Supplementary Figure 2. The first two columns were comprehensive report of Ramachandran plots of experimental structure 2WWX, 280K-320K replicate 1, and AlphaFold3 prediction replicate 1. Very few residues were in disallowed region, which revealed good standing in construction. The last column was the detailed report of hydrogen bonds calculated by PDBsum, related to residue 33-44 region of Rab1.



Supplementary Figure 3. The AlphaFold3 predication result of replicate 1. (a) most predicted regions are in very high (plDDT $>$ 90), confident (90 $>$ plDDT $>$ 70) levels, with few regions in low (70 $>$ plDDT $>$ 50) (b) Residue alignment and position error scores matrix; predication are in low bias.

Type	Interface area, Å^2	ΔG kcal/mol	No. of Hydrogen Bonds	No. of Salt Bridges	No. of non-bounded contacts
2WWX	1693.6	-13.4	24	3	247
AF rep0	1603	-8.8	23	4	206
AF rep1	1744.4	-12.6	22	3	228
AF rep2	1811.8	-13.3	22	2	215
320K rep1	1691.6	-13.7	26	2	228
300K rep1	1607.7	-9.8	22	2	188
280K rep1	1619.5	-8.8	20	4	196
320K rep2	1703.4	-12.3	23	2	203
300K rep2	1607.3	-10.9	22	2	167
280K rep2	1603.3	-9.1	20	4	193
320K rep3	1673.5	-13.1	25	3	221
300K rep3	1570.4	-10.4	19	4	172
280K rep3	1601.4	-8.6	22	4	170

Supplementary Table 1. Table of the detailed statistical indicators of all replicates. This table is the data references of figure 2b.