

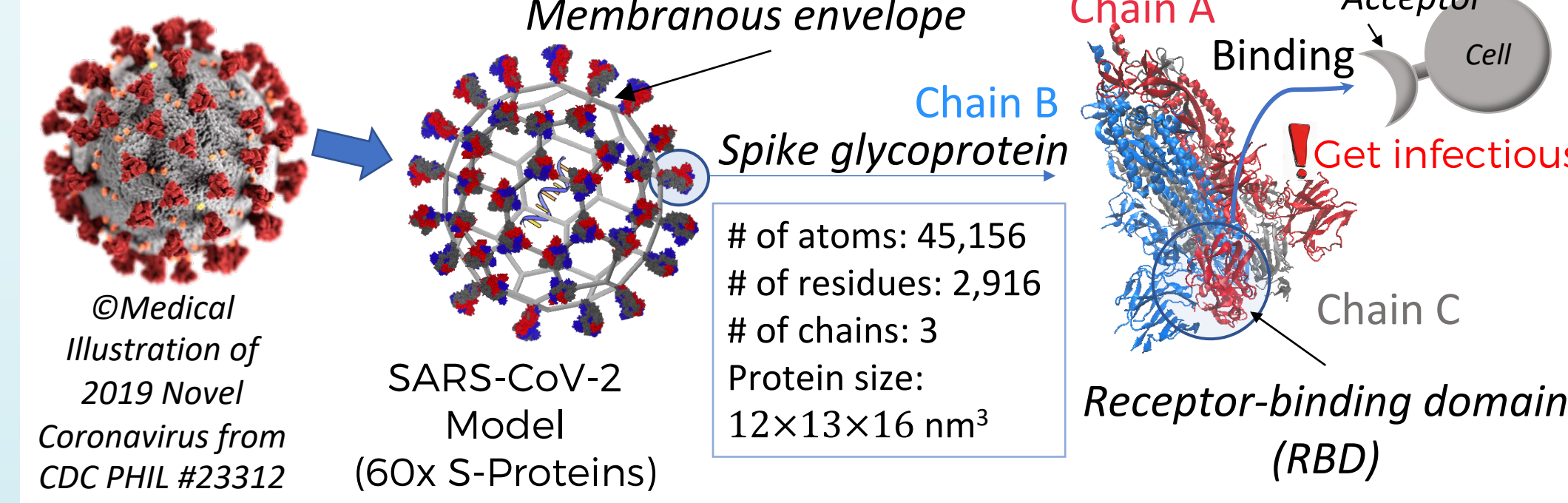
Long-Time Simulation of Temperature-Varying Conformations of SARS-CoV-2 Spike Glycoprotein on IBM Supercomputers

Meichen Song, Peng Zhang, Changnian Han, Ziji Zhang, Yuefan Deng

Applied Mathematics and Statistics Department, Stony Brook University, Stony Brook, NY, 11794, USA

Introduction & Motivation

A novel COVID-19 outbreak and spread worldwide in the past few months. The outer membrane spike glycoprotein (S-protein) is the prime host interacting protein with host cell targets (such as ACE2) important for cell adhesion and virulence [1]. In this study, we focus on the molecular dynamics (MD) modeling of the temperature-dependent properties of the S-protein, to offer a direct understanding of its conformational variations, and the possible critical temperature value.



Methodology

2- μ s MD Simulation of S-Protein Solvated in the Water Box

Structure Determines Function: Force Field

$$U_{MD} = \sum_{i \neq j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i \neq j} \frac{Q_i Q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{\text{bonds}} k_b (r - r_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{dihedrals}} k_\phi (1 + \cos(n\phi - \phi_0))$$

Simulation Temperature: 150°C -300°C
Denaturation phases detected: 95°C, 80°C, 70°C, 60°C, 37°C, 20°C, 3°C

Life Duration: Weeks, Days, Minutes, Seconds

Temperatures selected to locate the critical temperature: 150°C -300°C

Normal human body temperature: 37°C
Typical room temperature: 20°C
Cold supply chain temperature: 3°C

Structural Analysis

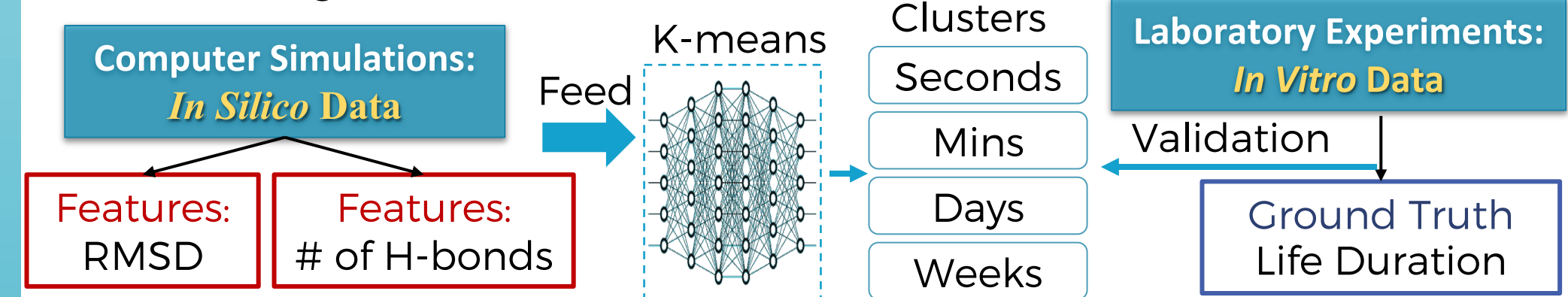
Protein Alignment: Center + Rotation

Time-dependent motions of the structure based on H-bonds and RMSD: whether a structure reaches the stable state and when the protein denature occurs?

Fluctuations of each subset of the protein based on RMSF: fluctuations of each residue relative to the average structure.

Minimize RMSD: 50ns simulation, Reference

Clustering for Life duration



HPC helps us to answer the question: Will the SARS-CoV-2 be more or less infectious as the ambient temperature increases?

The HPC Resources

AiMOS Supercomputer @ Rensselaer Polytechnic Institute

- IBM Power9 + NVIDIA V100 GPUs
- SeaWulf @ Stony Brook University

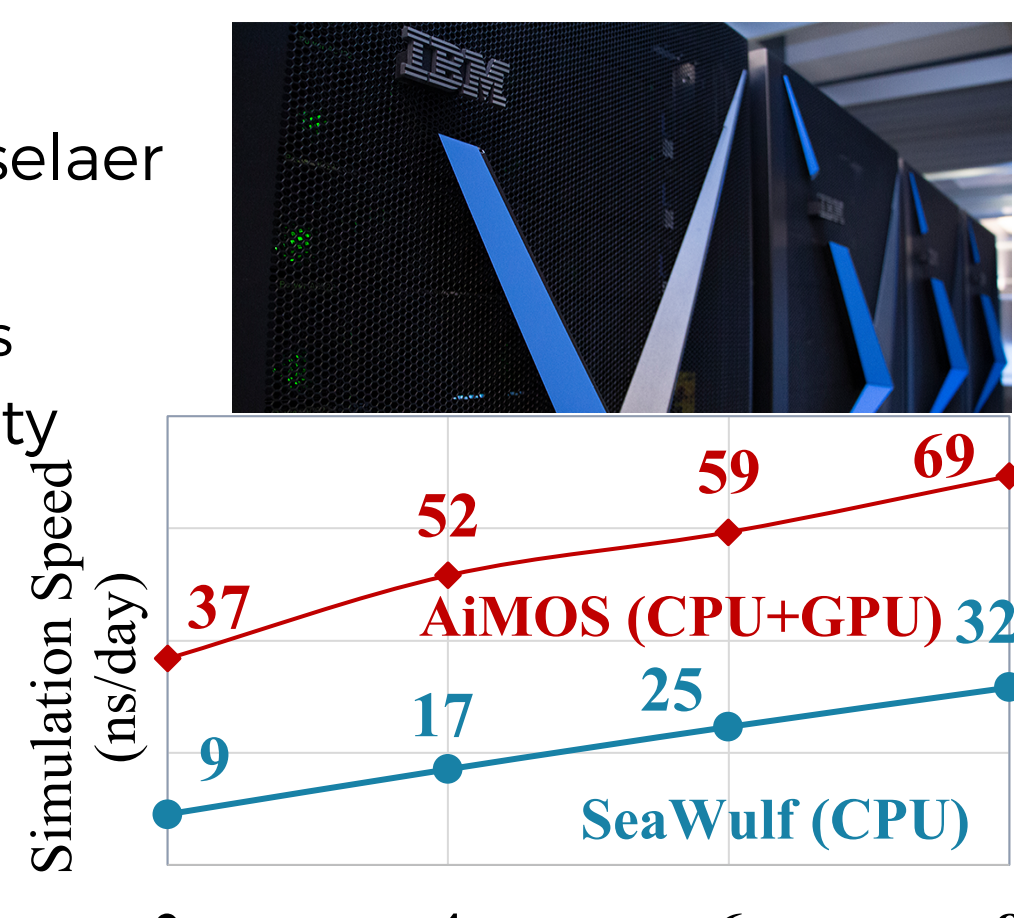
- Intel Xeon E5 CPUs
- HPC Performance

- A speedup of 7x for AiMOS over SeaWulf

- Reducing the 1-year task of a cluster to 2 months

- Enabled long-time MD simulation

AiMOS Information	POWER 9	V100
Number of CPUs/GPUs per node	2	6
Memory per node	2 x 256 GB	6 x 32 GB
Aggregate double-precision TFlop/s	0.97	40.3



Results: SARS-CoV-2 undergoes a temperature-mediated structural change at 60-80°C.

Time-dependent motions of the structure at temperatures below 100°C

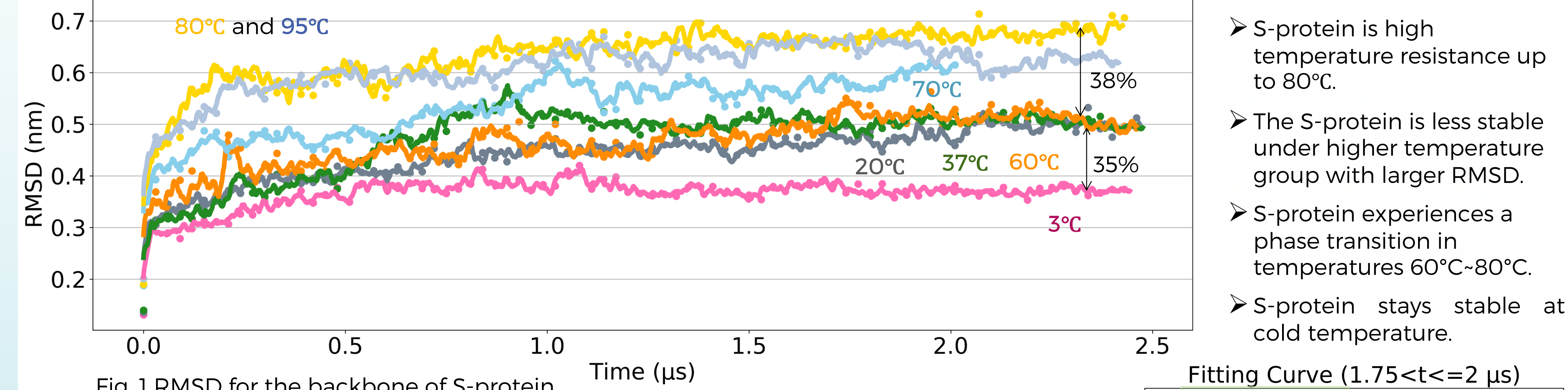


Fig. 1 RMSD for the backbone of S-protein

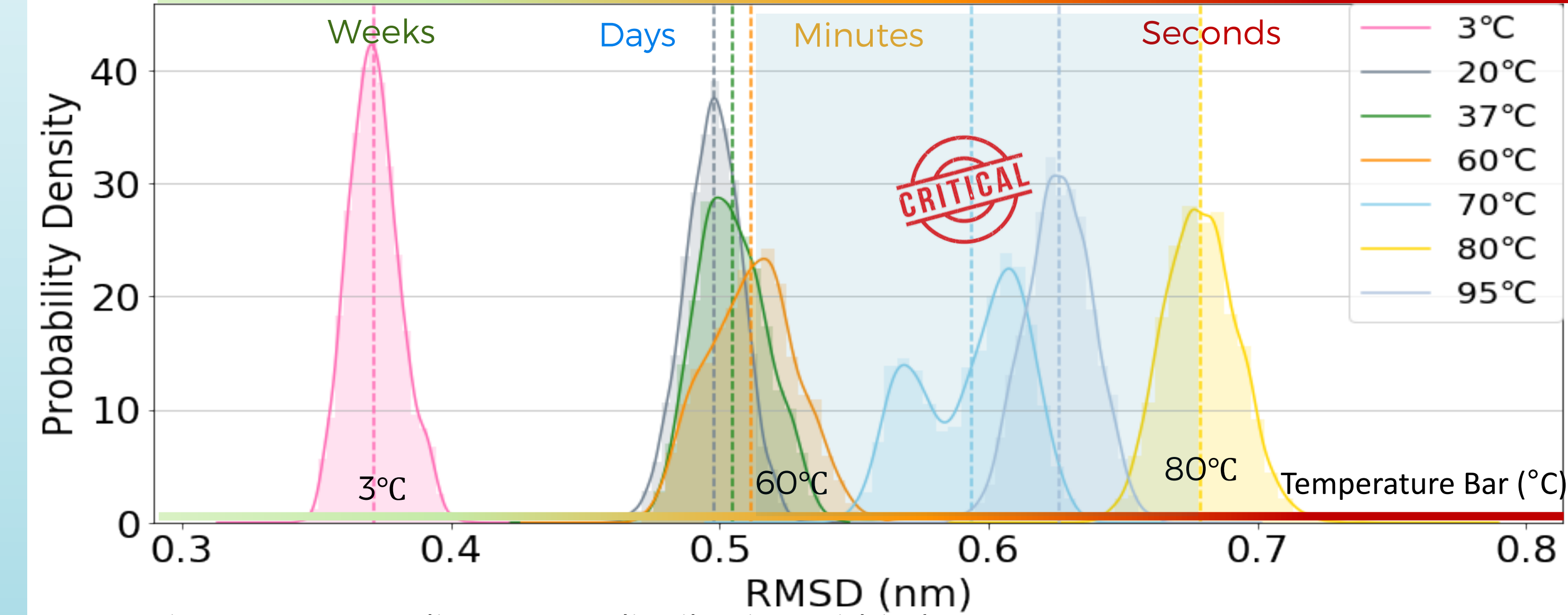
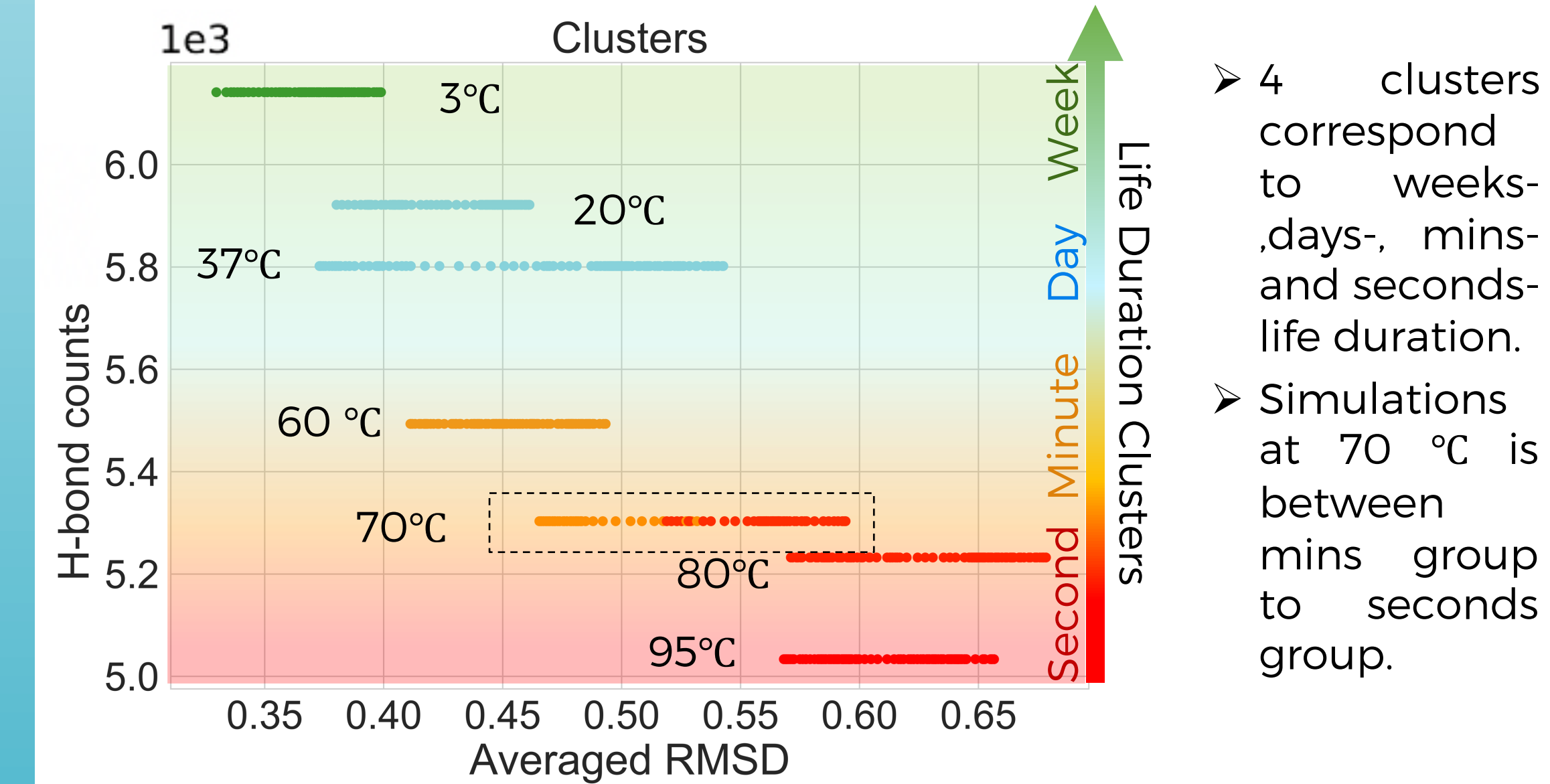
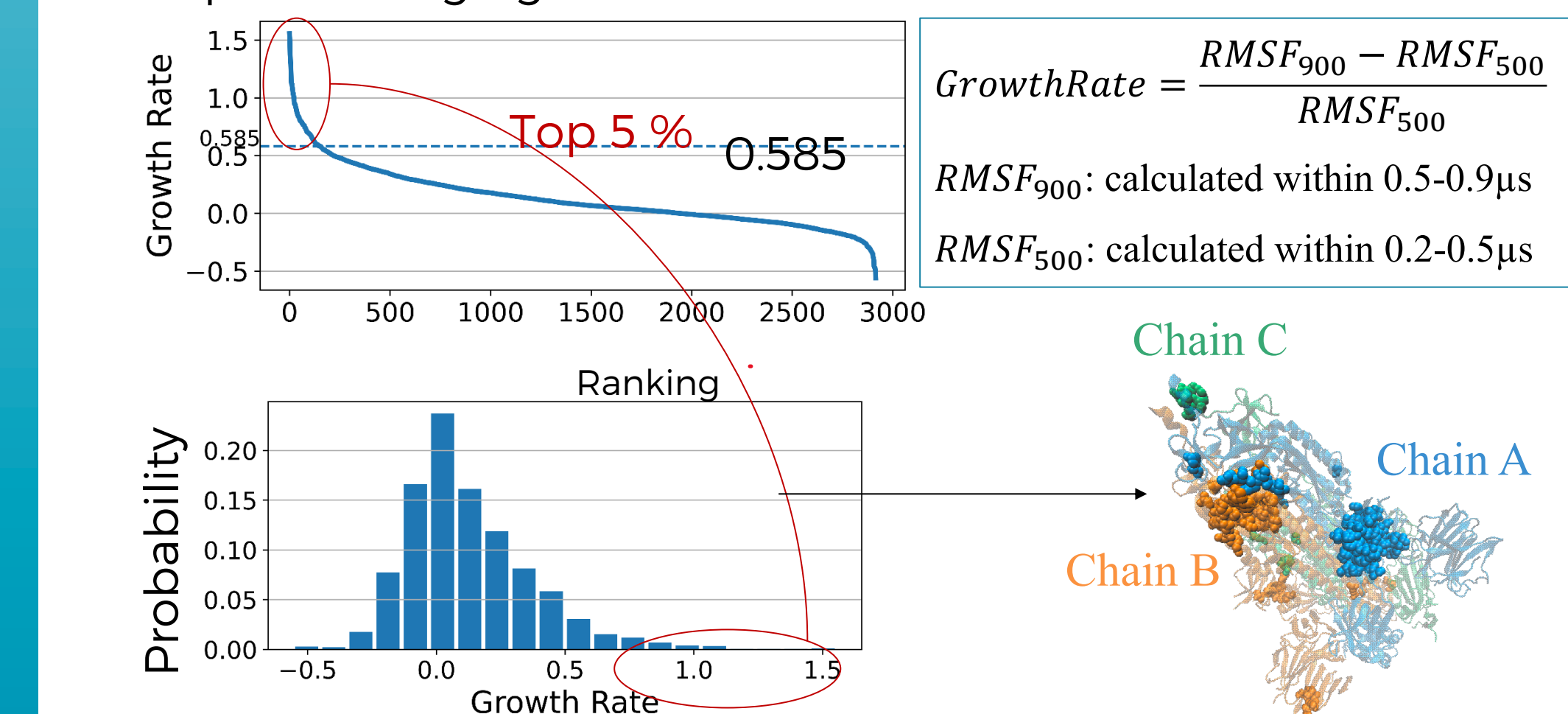


Fig. 2 Corresponding RMSD distribution within last 0.25 μ s

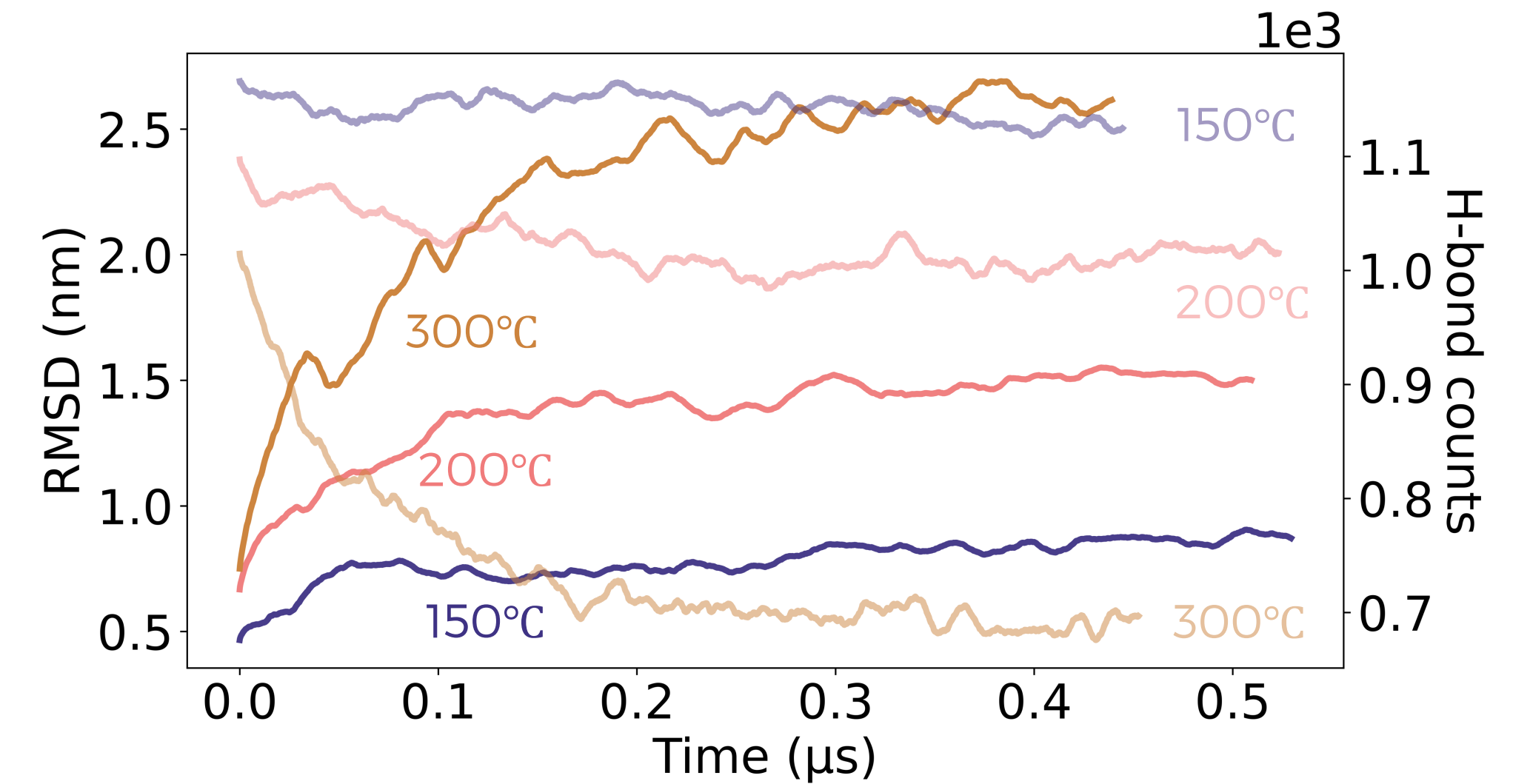
We introduce # of protein-water H-bonds to interpret the thermal instability of the S-protein from another perspective and cluster temperatures by K-means in terms of their life duration based on 6 features captured from RMSD and # of H-bonds (constants for each temperature). The clusters agree with lab-bench experiments [2].



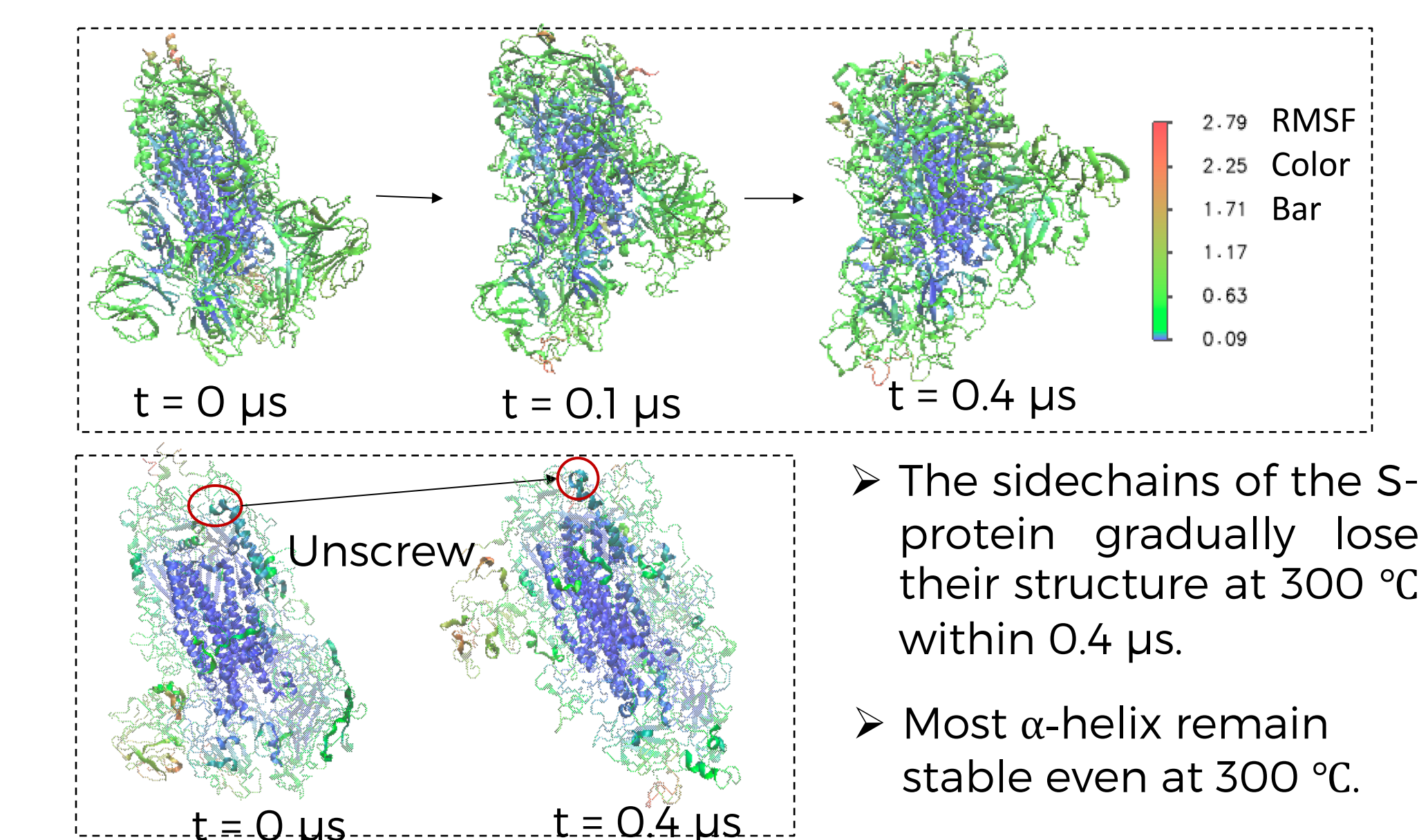
We define a measure of the growth rate based on RMSF to quantify the flexibility of residues on S-protein. We demonstrate the method with the case under the temperature 37°C and visualize residues of top five percent high growth rate.



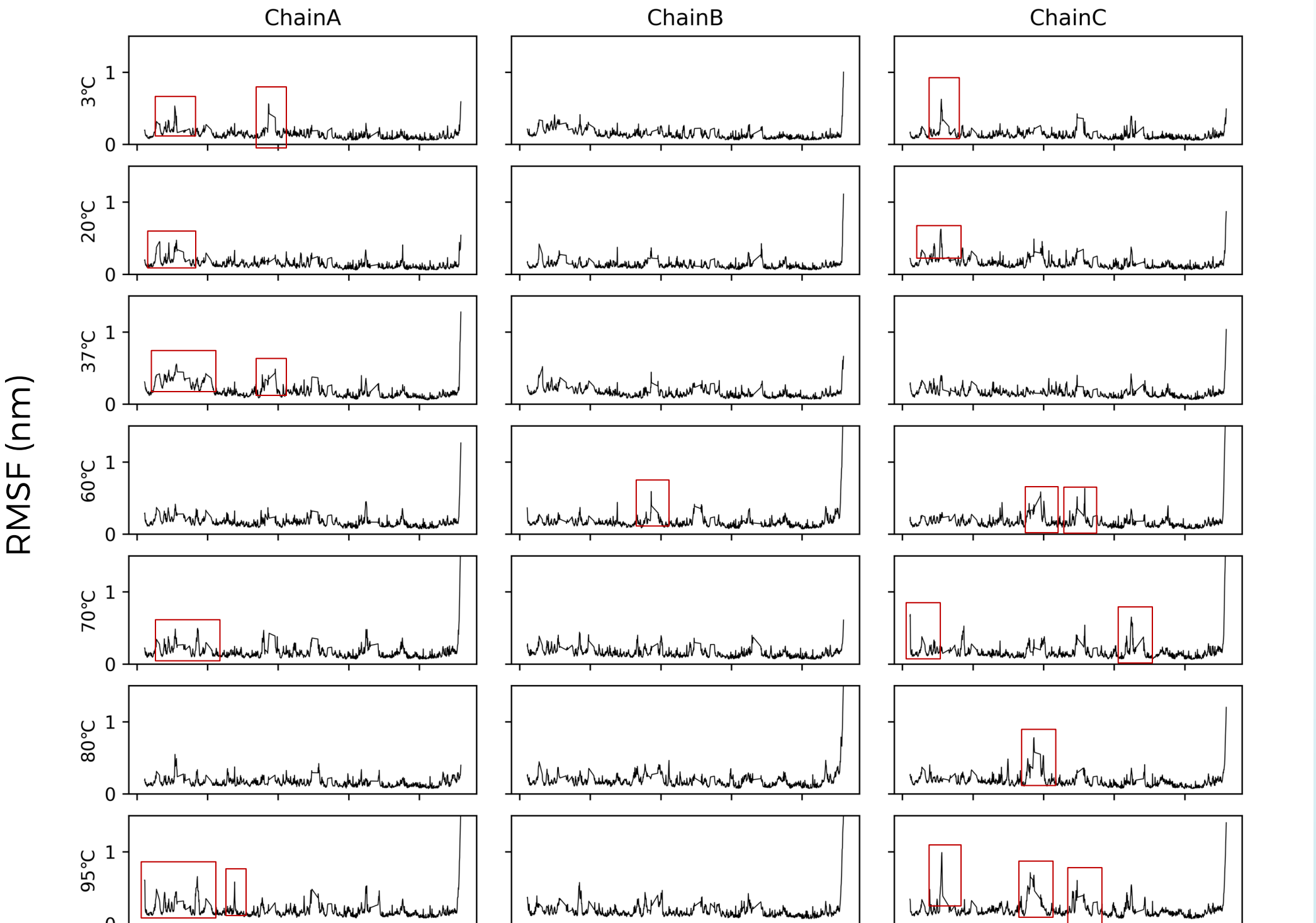
Time-dependent motions of the structure at temperatures higher than 100°C (which are ongoing running): RMSD for the backbone of S-protein (represented as darker lines) and mainchain-mainchain H-bonds (represented as lighter lines).



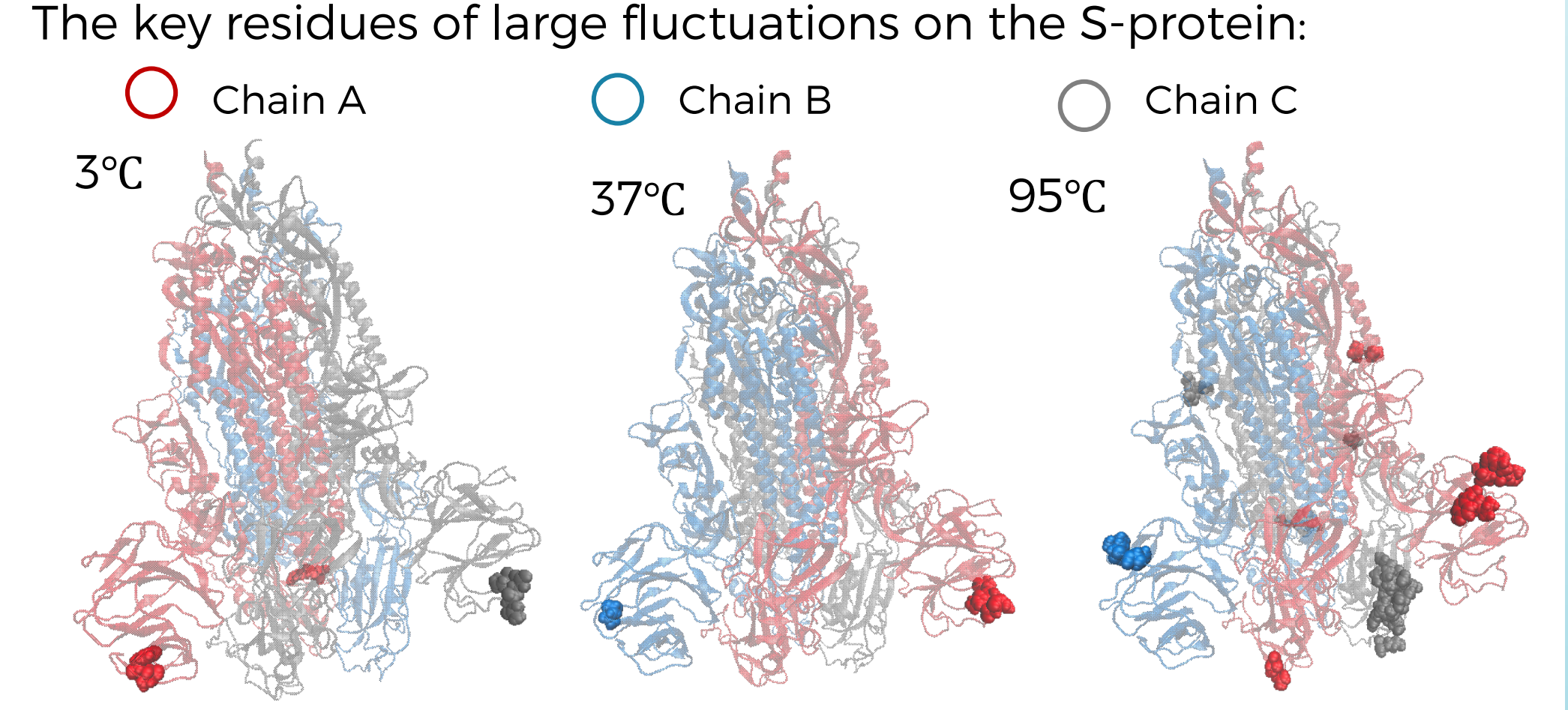
MD-HPC enable us to virtualize structure trajectory at 300°C to find out atom-level details for denaturation.



Residue-Based Analysis within 1.5-2.0 μ s.



The key residues of large fluctuations on the S-protein:

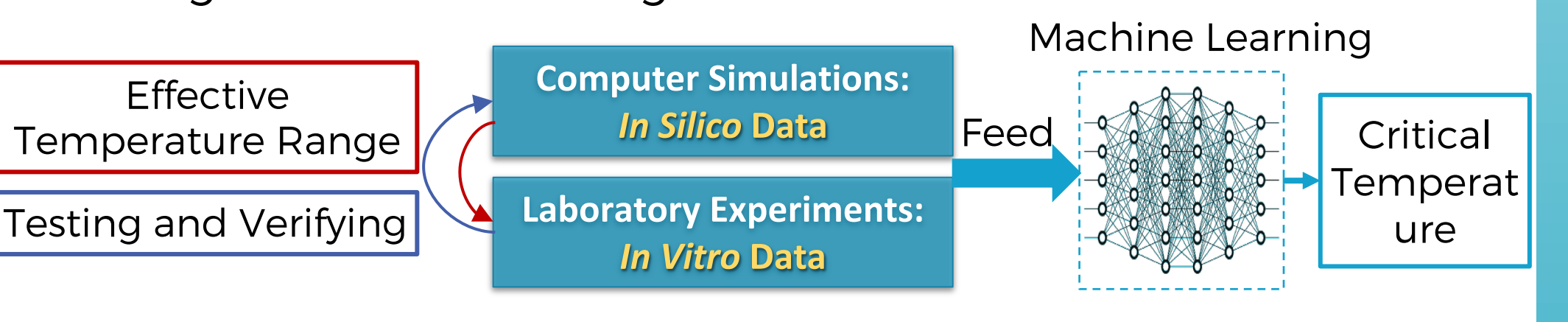


Conclusions & Future Work

The world's top supercomputers enable long-time MD simulations, filling the gap of our understanding of the temperature-mediated structural variation of the SARS-CoV-2 S-protein and enable us to explore the molecular-scale information, including key temperature-sensitive residues.

We will:

- Continue running simulation at three extremely high temperatures to explore the phases of denaturation for S-protein;
- Analysis more properties including the Gibbs free energy;
- Provide an effective temperature range for identifying the lethal temperature for the SARS-CoV-2 in vitro prediction through machine learning.



Acknowledgements & References

- The project is sponsored by the OVR&IEM COVID-19 Seed Grant, Pls: P. Zhang, Y. Deng, M. Rafailovich, M. Simon.
- The project is supported by the SUNY-IBM Consortium Award, Pl: Y. Deng, Co-I: P. Zhang.
- Core simulations were conducted on the AiMOS supercomputer at RPI and the WSC Cluster at IBM through an IBM Faculty Award FPO002468 (PI: Y. Deng).

[1] Song WF, Gui M, Wang X, et al. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with ACE2. PLoS Pathog. 2018;14(8):e1007236.
[2] A. W. Chin et al., "Stability of SARS-CoV-2 in different environmental conditions," 2020.