Dean&Francis
ISSN 2959-6157

Statistical Analysis of Nanobody Sequences and Structures with Molecular Dynamics Simulation of Nanobody VHH

Rui Xiao

Abstract:

Nanobodies, also known as single-domain antibodies, are an area of study due to their unique structure and small size. These antibodies, derived from camelids, consist only of heavy-chain domains, making them an excellent candidate for various therapeutic and diagnostic applications. This paper presents a statistical analysis of nanobody sequences and structures from existing databases and utilizes molecular dynamics (MD) simulation to model and assess the stability of a SARS-CoV-1 cross-reactive nanobody (6waq). By analyzing key metrics such as Root-Mean-Square Deviation (RMSD) and Root-Mean-Square Fluctuation (RMSF), this study provides insights into nanobody structural dynamics and potential optimizations for therapeutic use.

Keywords: nanobody, protein modelling, molecular dynamics, sequence analysis

1. Introduction

Nanobodies are a type of antibodies that have garnered significant attention due to their small size, stability, and ability to bind with antigens in a manner distinct from conventional antibodies. Discovered in camelids, these heavy-chain-only antibodies (HcAbs) consist solely of heavy-chain variable domains (VHH), making them smaller and more versatile than traditional antibodies. Due to their ability to penetrate molecular pockets and bind to epitopes that are otherwise inaccessible to larger antibodies, nanobodies have become a focal point for pharmaceutical research and biotechnology.

This study focuses on the statistical analysis of nanobody sequences and structures from the Protein Data Bank (PDB) and investigates the molecular dynamics of nanobody 6waq, which binds to the receptor-binding domain (RBD) of the SARS-CoV-1 spike protein. The study aims to understand the structural properties that contribute to the high stability and binding efficiency of nanobodies.

2. Experiments

2.1 Data Collection and Sequence Statistics

The data collection process began by gathering data from the Sab-nano antibody structure database, with a specific focus on nanobodies with existing sequence and structural information in the PDB. A series of custom Python scripts was developed to extract PDB IDs, retrieve amino acid sequences, and

ISSN 2959-6157

compile frequency statistics. This allowed for the analysis of the prevalence of certain amino acid residues and structural motifs specific to nanobodies.

2.2 Protein Modeling and Structural Evaluation

For the 6waq nanobody, the PDB structure was retrieved, and molecular modeling was performed using two platforms: SWISS-MODEL and AlphaFold2. Given the limitations of AlphaFold2 with multi-domain proteins, separate models were generated for chains A, B, C, and D. The models were evaluated using UCLA-DOE LAB — SAVES v6.0, with particular emphasis on structural stability, solvent accessibility, and stereochemistry. Hydrophobic interactions, hydrogen bonds, and salt bridges were predicted using ProteinTools.

2.3 Molecular Dynamics Simulation

Molecular dynamics simulations were conducted using the Amber suite. The A chain of 6waq was extracted, and a water solvent box was established using Charmm-Gui, maintaining a distance of 15 Å between the protein and the box boundary. NaCl was introduced to neutralize the system, and periodic boundary conditions (PBC) were applied. The system, comprising 57,476 atoms, underwent molecular dynamics simulation to analyze structural stability over time.

2.4 Structural Analysis

Key metrics analyzed during MD simulation included Root-Mean-Square Deviation (RMSD) and Root-Mean-Square Fluctuation (RMSF). These metrics provided insight into how the protein's backbone and individual residues shifted relative to the average structure over the course of the simulation.

3. Results

3.1 Sequence Statistics

The analysis suggested that nanobodies possess distinctive sequence motifs, with a higher prevalence of cysteine residues in complementarity-determining regions (CDRs), particularly in CDR1 and CDR3. These cysteines contribute to the formation of stabilizing disulfide bonds, which are essential for nanobody structural integrity under extreme conditions.

3.2 Structural Evaluation

Structural models generated using AlphaFold2 and SWISS-MODEL were compared, with both models showing high structural similarity. The Ramachandran plot analysis confirmed good stereochemical quality for both models, with over 95% of residues in favored regions. Hydrophobic interactions and hydrogen bonds were key stabilizing forces, especially in regions involved in antigen binding.

3.3 Molecular Dynamics Simulation

The MD simulation revealed that 6waq remained structurally stable over the course of the simulation. The RMSD values stabilized after 10 ns, suggesting that the nanobody reached equilibrium quickly. The RMSF analysis showed high fluctuations at the termini but relatively low fluctuation in the CDR regions, particularly CDR3, highlighting its role in maintaining antigen-binding stability. The residues Gly9, Gly55, and Gly98 exhibited the highest RMSF values, indicating areas of flexibility.

4. Discussion

The sequence and structural analysis confirm that nanobodies, owing to their compact size and stable framework, have a unique set of features that enable them to function effectively in diverse biological environments. The presence of disulfide bonds, particularly in the CDR regions, is a key contributor to their resilience against thermal and chemical denaturation.

The molecular dynamics simulation supports these observations, showing that nanobodies, like 6waq, maintain structural integrity even under physiological conditions. The flexibility observed in certain glycine residues might be critical for conformational adjustments during antigen binding. Moreover, the low RMSD values suggest that the nanobody's antigen-binding region remains stable, which is crucial for therapeutic applications where consistent binding is essential.

5. Conclusion

Nanobodies represent a promising tool in both therapeutic and diagnostic applications due to their small size, stability, and unique ability to target antigens in regions, which are inaccessible to conventional antibodies. Through sequence analysis and molecular dynamics simulation, this study highlights the structural stability and flexibility of nanobodies, offering insights into their binding efficiency and resilience. Future studies could further explore the role of flexible residues in antigen interaction and develop strategies to enhance nanobody circulation time in vivo.

6. References

[1] Hamers-Casterman, C. et al. (1993). Naturally occurring antibodies devoid of light chains. Nature.
[2] Muyldermans, S. (2001). Single domain camel antibodies: current status. Journal of Biotechnology.
[3] Kapila, C. et al. (2016). Molecular dynamics simulation of nanobodies. Journal of Computational Chemistry.