MECHANICAL, THERMAL AND TRANSPORT PROPERTIES OF MUTATED HEMOGLOBIN PROTEIN OF SICKLE CELL

A THESIS SUBMITTED TO THE

CENTRAL DEPARTMENT OF PHYSICS

INSTITUTE OF SCIENCE AND TECHNOLOGY

TRIBHUVAN UNIVERSITY

NEPAL

FOR THE AWARD OF

DOCTOR OF PHILOSOPHY

IN PHYSICS

BY

JHULAN POWREL

DECEMBER 2022

DECLARATION

Thesis entitled **"Mechanical, Thermal and Transport Properties of Mutated Hemoglobin Protein of Sickle Cell"** which is being submitted to the Central Department of Physics, Institute of Science and Technology (IOST), Tribhuvan University, Nepal for the award of the degree of Doctor of Philosophy (Ph.D.) is a research work carried out by me under the supervision of Prof. Dr. Narayan Prasad Adhikari of Central Department of Physics, Tribhuvan University, Nepal.

This research is original and has not been submitted earlier in part or full in this or any other form to any university or institute, here or elsewhere, for the award of any degree.

Jhulan Powrel

RECOMMENDATION

This is to recommend that **Mr. Jhulan Powrel** has carried out research entitled **"Mechanical, Thermal and Transport Properties of Mutated Hemoglobin Protein of Sickle Cell"** for the award of Doctor of Philosophy (Ph.D.) in **Physics** under my supervision. To my knowledge, this work has not been submitted for any other degree.

He has fulfilled all the requirements laid down by the Institute of Science and Technology (IOST), Tribhuvan University, Kirtipur for the submission of the thesis for the award of Ph.D. degree.

. . . . . .

Dr. Narayan Prasad Adhikari
Supervisor
(Professor)
Central Department of Physics
Tribhuvan University
Kirtipur, Kathmandu, Nepal

[December 2022]

LETTER OF APPROVAL

[Date: 19 /12/2022]

On the recommendation of **Prof. Dr. Narayan Prasad Adhikari**, this Ph.D. thesis submitted by **Mr. Jhulan Powrel**, entitled **"Mechanical, Thermal and Transport Properties of Mutated Hemoglobin Protein of Sickle Cell"** is forwarded by Central Department of Research Committee (CDRC) to the Dean, IOST, T. U.

. . . . . .

Dr. Om Prakash Niraula
Professor
Head
Central Department of Physics,
Tribhuvan University
Kirtipur, Kathmandu
Nepal

ACKNOWLEDGMENTS

I would like to express my hearty gratitude to supervisor Prof. Dr. Narayan Prasad Adhikari for guidance throughout my Ph.D. research work. As a supervisor, he has provided endless effort in the completion of this work.

I also acknowledge Prof. Dr. Om Prakash Niraula, HoD, Central Department of Physics for his official support to complete this thesis. I am thankful to Dean, IOST, T.U., Prof. Dr. Binil Aryal, who has always supported and motivated me in my research work.

I am greatly thankful to the CDRC members of Central Department of Physics. I would also thank the Dr. Dinesh Bhuju, Dr. Shankar Khanal and Dr. Chitra Bahadur Baniya for philosophical and statistical ideas in research.

I would also thank the teaching and non-teaching staffs of Central Department of Physics, T. U., Kirtipur. My sincere thanks goes to Dr. Gopi Chandra Kafle, Dr. Nurapati Pantha, Dr. Rajendra Prasad Koirala, Dr. Shyam Prakash Khanal, Dr. Saran Lamichane, Dr. Harikrishna Neupane for their kind suggestions and co-operation.

I am also thank for my lab mates Bidhya Thapa, Prakash Khatri, Narayan Gautam, Dharma Poudel. I am also thankful to all my friends especially Dr. Khimananda Neupane, Mr. Loknath Upadhyaya, Mr. Gopal Acharya, Mr. Indra Bhusal, Mr. Pradip Aryal, Mr. Laxman Gautam and Mr. Prakash Ghimire for their constant support throughout my Ph. D. study. At last my thank full words goes to my parents, Maya Poudel and Late Krishna Poudel for their limitless support throughout my course. At last thanks goes to Sumitra Neupane (Poudel), Abhinab Poudel and Angel Poudel and all other family members, relatives and well wisher for their motivating words towards my Ph.D. study.

Accordingly, I would like to acknowledge the University Grants Commission, Nepal for the Ph.D. Fellowship 76/77 of Science and Technology-13. I am equally thankful to Butwal Multiple Campus family for their great assistance. I am equally thankful to the New Horizon College fpr partial financial support.

Jhulan Powrel

December 2022

ABSTRACT

The Asia, Africa and states of the northern hemisphere are continuously suffering from sickle cell disease. Many of the physical factors have been studied interestingly for sickling of hemoglobin protein in the blood of sickle cell anemia patients. In the present work, we carry on a systematic study of sickle and non-sickle hemoglobin proteins to understand their structural, thermodynamics and transport properties at 310 K using the molecular dynamics (MD) technique. We have used the TIP3P water model as a solvent and all-atom charmm36m modified force field parameters to model our system.

We have estimated the root mean square deviation (RMSD), number of hydrogen bonds, salt bridges, hydrophobic, van der Waals (vdW) and electrostatic interactions and solvent accessible surface area (SASA) of sickle and normal hemoglobin proteins to investigate the structural conformation. We observed a higher number of hydrogen bonds, salt bridges, and hydrophobic interactions in sickle hemoglobin protein than that of normal. Also, sickle protein has shown a large value of vdW and electrostatics interactions in comparison to normal Hb protein tetramer. Our investigation shows that the SASA of normal hemoglobin is much less than that of sickle hemoglobin as expected in both tetramer and dimer, which may be due to the sickle shape of hemoglobin. The reduction in the contact area of the alpha chain indicates less bonding energy in the alpha chain with the other three chains in sickle hemoglobin protein than in the alpha of normal hemoglobin protein. This means sickle hemoglobin is more hydrophilic than normal hemoglobin. This finding indicates stronger confinement in sickle hemoglobin protein.

This work has further extended on identifying the binding components to realize the stiffness and their folding pathways in the sickle and normal hemoglobin proteins to explore the elastic properties using steered molecular dynamics (SMD). This study of beta and alpha chains in the hemoglobin proteins assures that a higher amount of force is required to separate the alpha and beta chain of normal hemoglobin than that of sickle hemoglobin. It also implies that the beta chain contributes more stiffness to the hemoglobin protein. The force for breaking the bonding in alpha chain in normal protein is higher than in sickle indicating that the alpha chain has higher stiffness in normal than in sickle hemoglobin protein. Both studies indicate higher stiffness in sickle hemoglobin is due to the contribution of the beta chain. It supports the theoretical concept of higher free energy in normal than in sickle hemoglobin protein.

Moreover, the specific heat capacity of normal hemoglobin and sickle hemoglobin proteins have been estimated. The normal hemoglobin protein has a higher specific heat capacity than that of the sickle hemoglobin protein. Also, the self and binary diffusion

coefficients of sickle is obtained less than that of normal hemoglobin protein dimer. The diffusion coefficient of sickle and normal hemoglobin are also increased with temperature in usual way. We have also estimated the binding free energy of dimerization. The binding free energy of alpha and beta chains in the hemoglobin protein dimer structure of sickle and normal is found to be (5.97±0.27) kcal/mol and (6.64±0.27) kcal/mol. The difference in free energy in the sickle and the normal hemoglobin is estimated as (0.67±0.06) kcal/mol. Thus, it is recommended to increase the free energy carried by sickle protein to work as normal protein.

LIST OF ACRONYMS AND ABBREVIATIONS

AMBER	:Assisted Model Building with Energy Refinement
BAR	:Bennett Acceptance Ratio
CHARMM	:Chemistry at Harvard Macromolecular Mechanics
Da	:Dalton
DNA	:Deoxyribo-Nucleic Acid
EM	:Energy Minimization
ER	:Excluded Region
FEP	:Free Energy Perturbation
fs	:Femtosecond
GROMACS	:Groningen Machine for Chemical Simulation
Hb	:Hemoglobin
LINCS	:LINear Constraint Solver
LJ	:Lennard Jones
MD	:Molecular Dynamics
MSD	:Mean Squared Displacement
NAMD	:NAnoscale Molecular Dynamics
NMR	:Nuclear Magnetic Resonance
NPT	:Constant Number, Pressure and Temperature
NVE	:Constant Number, Volume and Energy
NVT	:Constant Number, Volume and Temperature
OPLS-AA	:Optimized Potentials for Liquid Simulations-All Atom
PBC	:Periodic Boundary Conditions
PDB	:Protein Data Bank
PME	:Particle Mesh Ewald
RBC	:Red Blood Cell
RMSD	:Root Mean Square Deviation
RNA	:Ribonucleic Acid

SASA :Solvent Accessible Surface Area
SMD :Steered Molecular Dynamics
TIP3P :Transferable Intermolecular Potential with 3 Points
vdW :van der Waals
VMD :Visual Molecular Dynamics
WHAM :Weighted Histogram Analysis Method

LIST OF SYMBOLS

k_{B} :Boltzmann constant
k_{ij}^{b} :Force constant for bond stretching potential
b_{ij}^{0} :Equilibrium bond length
k_{ijk}^{θ} :Force constant for bond angle potential
θ_{ijk}^{0} :Equilibrium bond angle
k_{ϕ} :Force constant for dihedral angle potential
δ :Equilibrium dihedral angle in degree
k_{ξ} :Force constant for improper dihedral angle potential
ξ_0 :Equilibrium improper dihedral angle in degree
ϵ_0 :Permittivity of free space
ϵ_r :Dielectric constant
ϵ :Strength of potential in L-J potential
σ :Length scale in L-J potential
k_i :Force constant for i^{th} Window in umbrella sampling
α :Alpha chain in the hemoglobin protein
β :Beta chain of hemoglobin protein
D :Diffusion constant of Hb protein
k :Spring constant
c_v :Specific heat capacity at constant volume

LIST OF TABLES

Page No.

Table 1: The number of hydrophobic sets of sickle and normal hemoglobin dimer protein at five different consecutive time frames (Powrel et al., 2022). 61

Table 2: The average vdW and electrostatic interaction potential energy in sickle hemoglobin protein and its alpha chain (Powrel et al., 2022). 66

Table 3: The average vdW and electrostatic interaction potential energy in normal hemoglobin protein and its alpha chain (Powrel et al., 2022). 66

Table 4: The average vdW and electrostatic interaction potential energy in sickle hemoglobin protein and its beta chain (Powrel & Adhikari, 2022). 67

Table 5: The average vdW and electrostatic interaction potential energy in normal hemoglobin protein and its beta chain (Powrel & Adhikari, 2022). 67

Table 6: The averages of vdW and electrostatics interaction potential energy of sickle hemoglobin protein dimer. 67

Table 7: The averages of vdW and electrostatics interaction potential energy of normal hemoglobin protein dimer. 67

Table 8: SASA of HbS and A3N protein with k=5, k=8 and k=11 kcal mol $^{-1}$ Å^{-2}. 71

Table 9: SASA distribution with time of chain α (A), chain BCD and protein of both sickle and normal hemoglobin dimer in nm^2. . . 73

Table 10: SASA distribution with time of α chain, β chain andwhole proteins of both sickle and normal hemoglobin dimer in Å^2 (Powrel et al., 2022). 74

Table 11: Forces (pN) for H-bond breaking in sickle hemoglobin (HbS) and normal hemoglobin (A3N) with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022). 77

Table 12: Forces (pN) required for breaking the H-bonds in the alpha chain of sickle (HbS) and normal hemoglobin (A3N) protein with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2}. 83
Table 13: For normal and sickle cell hemoglobin protein at 310K of temperature(Powrel & Adhikari, 2021). 89
Table 14: MSD of sickle hemoglobin at temperatures of 306 K, 308 K, 310 K and 312 K. 91
Table 15: Slope of MSD graph of normal hemoglobin at temperature of 306 K, 308 K, 310 K and 312 K. 92
Table 16: Temperature dependency of diffusion Coefficient of TIP3P water, sickle and normal hemoglobin Proteins. 93

LIST OF FIGURES

Page No.

Figure 1: (a) RBC with sickle (left) and normal (right) shape and (b) impact of sickle cell dieasease. 3

Figure 2: Heme ligands of (a) sickle hemoglobin protein with valine and (b) normal hemoglobin with glutamic acid in 6^{th} position of beta chain. 5

Figure 3: (a) Binding site of dimer in HbS and (b) surface interaction in $\alpha - \beta$ chain dimer of sickle hemoglobin. 6

Figure 4: Dimer structure of (a) sickle and (b) normal hemoglobin proteins. 7

Figure 5: Hemoglobin protein with (a) double α-β dimer structure, (b) sickle hemoglobin and (c) normal hemoglobin protein. 9

Figure 6: First seven residues with (a) valine at 6th position of beta chain in sickle and (b) glutamic acid residue at 6th position of beta chain in normal hemoglobin. 9

Figure 7: Tetramere view of HbS molecule (Powrel et al., 2022). 10

Figure 8: (a) Bond stretching structure and (b) bond stretching potential. 25

Figure 9: (a) Bond angle structure and (b)bond angle potential. 26

Figure 10: (a) Improper dihedral structure and (b) improper dihedral potential. 26

Figure 11: a) Proper dihedral structure and b) periodic dihedral potential. 27

Figure 12: Lennard-Jones potential 28

Figure 13: (a) Representation of two-dimensional simulation boxes with molecules in a repeated way applying periodic boundary conditions and (b) consideration of cut-off for interaction and (c) Simulation box with protein molecule and TIP3P water. 30

Figure 14: Flow chart of steps followed in molecular dynamics simulation. 33

Figure 15: Snap shots of five consecutive frames of N-C terminal SMD of beta chain in sickle hemoglobin 34

Figure 16: Ten consecutive snapshot of USSMD of dimer structure in sickle hemoglobin (Powrel et al., 2022). 38

Figure 17: Software packages of (a) NAMD, (b) PyMOL (c) VMD and (d) GROMACS. 40

Figure 18: (a) Energy profile versus reaction coordinate and (b) probability distribution of biasing potential versus reaction coordinate. 43
Figure 19: Block diagram showing the relation of computer simulations with experiment and theoretical study. 44
Figure 20: Hemoglobin protein with (a) double - dimer structure, (b) sickle hemoglobin, (c) normal hemoglobin protein, (d) sickle hemoglobin dimer and (e) normal hemoglobin dimer. 47
Figure 21: (a) RMSD of sickle and normal hemoglobin (b) distribution of RMSD (Powrel & Adhikari, 2022). 48
Figure 22: RMSD of (a) sickle hemoglobin and (b) RMSD distribution. . . 48
Figure 23: (a) RMSD of sickle and normal hemoglobin (b) distribution of RMSD. 49
Figure 24: a) RMSD of sickle hemoglobin dimer and b) distribution of RMSD (Powrel et al., 2022) 50
Figure 25: a) RMSD of normal hemoglobin dimer and b) distribution of RMSD (Powrel et al., 2022) 50
Figure 26: (a) Time evolution of hydrogen bonds in sickle and normal hemoglobin protein and (b) distribution of the number of hydrogen bonds in sickle and normal hemoglobin protein tetramer (Powrel & Adhikari, 2022) 52
Figure 27: Average occupancy of H-bond versus pairs of residues involved in sickle and normal hemoglobin tetramer (color online) (Powrel & Adhikari, 2022). 52
Figure 28: (a) Number of hydrogen bonds in the sub-units and (b) distribution of the hydrogen bonds in sickle Hb tetramer. 53
Figure 29: (a) Number of hydrogen bonds in the sub-units and (b) distribution of the number of hydrogen bonds in normal Hb. 54
Figure 30: Average occupancy of hydrogen bond versus residue pairs involved in (a) sickle and (b) normal hemoglobin tetramer during alpha SMD. 55
Figure 31: (a) H-bond in sickle hemoglobin dimer and (b) average number of the H-bonds. 56
Figure 32: (a) Number of H-bond in normal hemoglobin dimer and (b) distribution of number of the H-bonds (Powrel et al., 2022)(color online). 57
Figure 33: Average occupancy of hydrogen bond versus residue pairs involved in (only above 40 %) sickle hemoglobin dimer (Powrel et al., 2022). 58

Figure 34: Average occupancy of hydrogen bond versus residue pairs involved in (only above 40 %) normal hemoglobin dimer (Powrel et al., 2022). 58
Figure 35: a) Hydrophobic nature of 6^{th} valine in in β chain of sickle and b) hydrophilic nature of 6^{th} glutamic residues in β chain of normal hemoglobin protein. 60
Figure 36: Graph of the average salt bridge occupancy of different residue pairs in the sickle and normal hemoglobin protein (color online). 62
Figure 37: Salt bridges in PROA of (a) sickle and (b) normal hemoglobin tetramer. 63
Figure 38: Salt bridge occupancy in sickle hemoglobin dimer (Powrel et al., 2022). 64
Figure 39: Salt bridge occupancy in normal hemoglobin dimer (Powrel et al., 2022). 65
Figure 40: The SASA (a) of beta chain (chain B) in sickle and normal hemoglobin tetramer and (b) SASA distribution with time (Powrel & Adhikari, 2022). 69
Figure 41: SASA (a) of A, C and D chain alone of sickle and normal hemoglobin protein and (b) SASA distribution with time (Powrel & Adhikari, 2022). 69
Figure 42: SASA (a) of sickle and normal hemoglobin protein tetramer and (b) SASA distribution. 70
Figure 43: (a) Plot of SASA of equilibrated beta chain in sickle protein complex (b) Plot of SASA of equilibrated beta chain in normal hemoglobin protein complex for for k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022). 71
Figure 44: SASA of a) sickle HbS protein and b) SASA distribution with time.. 72
Figure 45: SASA of a) normal Hb protein and b) SASA distribution with time. 73
Figure 46: (a) SASA's in sickle hemoglobin dimer and (b) SASA distribution with time (Powrel et al., 2022) 74
Figure 47: (a) SASA's in normal hemoglobin dimer and (b) SASA distribution with time (Powrel et al., 2022). 74
Figure 48: Snapshots of SMD of alpha chain of sickle hemoglobin protein. 75

Figure 49: SMD graph of the beta chain in HbS protein with a) k=5, b) k=8, c) k=11 kcal mol $^{-1}$ Å^{-2} and d) SMD graph of A3N protein with k=5 kcal mol $^{-1}$ Å^{-2} e) SMD graph of A3N protein with k=8 and f) SMD graph of A3N protein with k= 11 kcal mol $^{-1}$ Å^{-2} (Powrel & Adhikari, 2022). 76

Figure 50: a) Pulling velocity versus force required graph of HbS with k=5 kcal mol^{-1} Å^{-2} b) Pulling velocity versus force required graph of HbS with k=8 kcal mol^{-1} Å^{-2} kcal mol^{-1} Å^{-2} c) Pulling velocity versus force required graph of HbS with k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022). 78

Figure 51: a) Pulling velocity versus force required graph of A3N with k=5 kcal mol^{-1} Å^{-2} b) Pulling velocity versus force required graph of HbS with k=8 kcal mol^{-1} Å^{-2} and c) Pulling velocity versus force required graph of HbS with k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022). 79

Figure 52: Pulling velocity versus force required graph for beta chain of (a) HbS with different k with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2}.and (b) A3N with different with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022). 80

Figure 53: SMD graph of alpha chain in HbS protein with a) k=5, b) k=8, c) k= 11 kcal mol $^{-1}$ Å^{-2}, d) SMD graph of alpha chain in A3N protein with k=5 kcal mol $^{-1}$ Å^{-2}, e) SMD graph of A3N protein with k=8 and f) SMD graph of A3N protein with k= 11 kcal mol $^{-1}$ Å^{-2}. 82

Figure 54: (a) Pulling velocity versus force required graph of HbS with K=500 (b) Pulling velocity versus force required graph of HbS with K=800 kcal mol^{-1}nm^{-2} and (c) Pulling velocity versus force required graph of HbS with K=1100 kcal mol^{-1}nm^{-2} 84

Figure 55: (a) Pulling velocity versus force required graph of A3N with k=500 (b) Pulling velocity versus force required graph of A3N with k=800 kcal mol^{-1}nm^{-2} and (c) Pulling velocity versus force required graph of A3N with k=1100 kcal mol^{-1}nm^{-2} 85

Figure 56: Pulling velocity versus force required graph for alpha chain SMD at different k of k=500 kcal mol^{-1} nm^{-2}, k=800 kcal mol^{-1} nm^{-2} and k=1100 kcal mol^{-1} nm^{-2} for (a) HbS protein and (b) A3N protein. 86

Figure 57: (a) Total energy fluctuation of sickle and normal hemoglobin protein at 310 K and (b) distribution of energy (Powrel & Adhikari, 2021). 88
Figure 58: MSD versus time plots of sickle hemoglobin at (a) 306 K, (b) 308 K, (c) 310 K and (d) 312 K. 90
Figure 59: MSD versus time plots of normal hemoglobin at (a) 306 K, (b) 308 K, (c) 310 K and (d) 312 K. 91
Figure 60: MSD versus time plots of water with hemoglobin protein at (a) 306 K, (b) 308 K, (c) 310 K and (d) 312 K.. 92
Figure 61: Binding in (a) sickle HbS and (b) normal hemoglobin dimer (Powrel et al., 2022). 94
Figure 62: Free energy graph of sickle and normal hemoglobin proteins (Powrel et al., 2022) 94

TABLE OF CONTENTS

Page No.

Declaration i
Recommendation ii
Letter of Approval iii
Acknowledgements iv
Abstract v
List of Acronyms and Abbreviations vii
List of Symbols ix
List of Tables x
List of Figures xii

CHAPTER 1 1

1. INTRODUCTION 1
1.1 General Consideration 1
1.1.1 Dimer Structure of Sickle and Normal Hemoglobin Proteins . . 6
1.1.2 Tetramer Structure of Sickle and Normal Hemoglobin Protein . 8
1.1.3 Importance of Steered Molecular Dynamics (SMD) 9
1.1.4 Binding Free Energy Difference 10
1.1.5 Thermal and Transport Properties 11
1.2 Rational of the Study 13
1.3 Objectives of the Study 14
1.4 Organization of the Thesis 15

CHAPTER 2 16

2. LITERATURE REVIEW 16

2.1 Genetral Consideration 16

CHAPTER 3 21

3. MATERIALS AND METHODS 21

3.1 General Consideration 21

3.2 Theoretical Background 21

3.2.1 Free Energy 22

3.2.2 Specific Heat Capacity 23

3.2.3 Diffusion 23

3.3 Mpolecular Dynamics 24

3.4 Computational Details 35

3.4.1 Simulation Details 37

CHAPTER 4 46

4. RESULTS AND DISCUSSION 46

4.1 General Consideration 46

4.2 Results and Discussion 46

4.2.1 Root Mean Square Deviation 47

4.2.2 Hydrogen Bond Interactions 51

4.2.3 Hydrophobic Interactions 59

4.2.4 Salt Bridge Interactions 61

4.2.5 vdW and Electrostatic Interactions 65

4.2.6 Solvent Accessible Surface Area (SASA) Calculations 68

4.2.7 Elastic Property by Steered Molecular Dynamics 75

4.2.8 Estimation of Specific Heat Capacity 88

4.2.9 Diffusion of Sickle and Normal Hemoglobin Dimer 89

4.2.10 Binding Free Energy by Umbrella Sampling 93

CHAPTER 5 96

5. CONCLUSIONS AND RECOMMENDATIONS 96

5.1 General Consideration 96

5.1.1 Elastic Property of Sickle and Normal Hemoglobin Tetramer Protein 96

5.1.2 Free Energy of Sickle and Normal Hemoglobin Dimer Proteins 97

5.1.3 Specific Heat Capacity of Sickle and Normal Hemoglobin Proteins 98

5.1.4 Diffusion of Sickle and Normal Hemoglobin 98

CHAPTER 6 100

6. SUMMARY 100

REFERENCES 101

APPENDIX 118

A. Articles Published in International Journals

B. Articles published in National Journals

C. Articles in Review

D. Participation in Conferences, Seminar and Meeting

CHAPTER 1

1. INTRODUCTION

1.1 General Consideration

In living organism several genetic disorder occurs due to mutation in DNA sequence as well as in protein. Such a mutation occurs in the hemoglobin protein molecule of human being is sickle cell disease (Pauling et al., 1949). It has been investigated over the last fifty years. In this disease the hemoglobin protein takes sickle like shape instead of spherical due to single amino acid replacement in the normal hemoglobin (Ingram, 1956). Asia, Africa and the states of northern hemisphere are continuously suffering the problems of sickle cell diseases (Luzzatto, 2012). Dormandy et al. (2018) reported 14,000 people in the UK have an inherited sickle cell disorder which, mainly impacts the people of black African and African Caribbean heritage. A research of 2019 indicates that about 1250 sickle patients are still present in Western region of Nepal (Pande et al., 2019). The disease causes chronic fatigue, pain, swelling, damage to vital organs and leads to early mortality etc. A genetic disorder in the patient of sickle cell diseases seems health issue in the present day.

Hemoglobin is a transport material found in the human blood which maintains the body temperature and carries oxygen, carbon dioxide and other nutrient materials throughout the body. Heat, minerals and ions are also conducted through the blood. The red blood corpuscle (RBC) of blood is composed of hemoglobin protein. The globular hemoglobin protein is iron rich and is present in the red blood cells of vertebrates as well as in some invertebrates (Jia et al., 2016). Oxygen bubbles when combines with hemoglobin protein molecules in the lungs the blood becomes oxygenated and transported to the body part to keep them alive (M. F. Perutz, 1978; Krogh, 1941). The iron rich oxygenated hemoglobin gives the red color to the blood (Billett, 1990; Antonini & Brunori, 1970). Besides carrying oxygen and carbon monoxide, hemoglobin has other essential functions like, genetic resistance to malaria, heat transportation in the body through its oxygenation deoxygenation cycle and other enzymatic activities (Giardina et al., 1995). Hemoglobin protein is interactive with other three gases of carbon monoxide

(CO), carbon dioxide (CO_2), water vapor and nitric oxide (NO) (Schechter, 2008).

A normal hemoglobin (Hb) molecule has the molecular weight of 64,500 Da containing two α and two β poly peptide chains (Northrop & Anson, 1929). It is synthesized in a number of steps. The heme is iron containing unit synthesized in the mitochondria of cell, while the four globular proteins are synthesized through ribosomes in the cytosole. The heme ligand is surrounded by four globin protein for its protection which is vital for oxygen transport. These globin chains are interacted with each other by certain energy. This energy is the contribution of the non bonded energy of vdW, electrostatics, salt bridge, hydrogen bond, hydrophobic interaction etc.

Beta-globin is a sub-unit of a larger protein called hemoglobin, which is located inside red blood cells. The beta-globin is formed by the instruction of HBB genes at the bone marrow. The adult hemoglobin consists of four peptide sub units: two sub units of beta-sheet and two sub-units of a protein called alpha-helix, which is produced from another gene called HBA. Each of these protein sub units are attached to an iron containing molecule of heme; such a heme contains an iron molecule at its center which can bind to one oxygen molecule. The T-state (taut) hemoglobin is the deoxy state i.e., lacking of oxygen and R-state is the relaxed state by carrying oxygen in it (Monod et al., 1965). The two states differ in ligand binding affinity and interactions between the chains in the complex. The RBC contains hemoglobin protein, which are like tiny bubbles of spherical shape and combine with the air (oxygen) in the lungs. such oxygenated blood moves round our body parts to keep the body tissues and organs alive by carrying oxygen to them. Hemoglobin gives the red color to blood when it contains oxygen. These bubbles of hemoglobin molecules stay freely flowing in the red blood cell. There are different types of normal hemoglobin flowing in red blood cell; adult hemoglobin, fetal hemoglobin and minor adult hemoglobin. The types of hemoglobin gene inherited from their parents determines the type of adult hemoglobin (HbA) they will have. Genes always comes in pairs; one from the mother and another from the father. Therefore, every person inherits two adult hemoglobin genes. The normal and most common hemoglobin gene sequence that an individual can inherit from their parent is hemoglobin AA (Hb AA). Hemoglobin contains iron and protein which makes the red blood cells and helps in transport oxygen throughout the human body. Higher level of hemoglobin could be the polycythemia and mutation in the beta chain of hemoglobin is sickle cell (Pauling et al., 1949). It is structurally a globular protein with a spherical shape Figure 1 (a) (left) (Retrieved December 15, 2022 (https///www.google/Sickle Cell)) having primary, secondary and tertiary structure.

The peptide chain of alpha comprises with 141 and beta with 146 amino acid residues. The chain alpha with an atomic weight of 15.15 KDa and the chain beta with atomic weight of 15.86 KDa. In some of the inherited hemoglobin protein, the amino acid se-

quence are changed due to the incorrect deoxyribonucleic acid (DNA) code (HbS) (Ribeil et al., 2017). Sickle cell iron deficiency is one of the such variety of transformation in which there is substitution of adenine within the 6th arrangement of the beta quality (GAG-GTG), which encodes the valine buildup rather than glutamic acid within the 6th position of beta chain of sickle hemoglobin (Noguchi & Schechter, 1981; Shrestha & Karki, 2013). A hemoglobin cell having such abnormality appears sickle like shape. This polymerization results in the RBC losing the deformability properties of the RBC and taking a sickle like shape Figure 1 (a)(right).

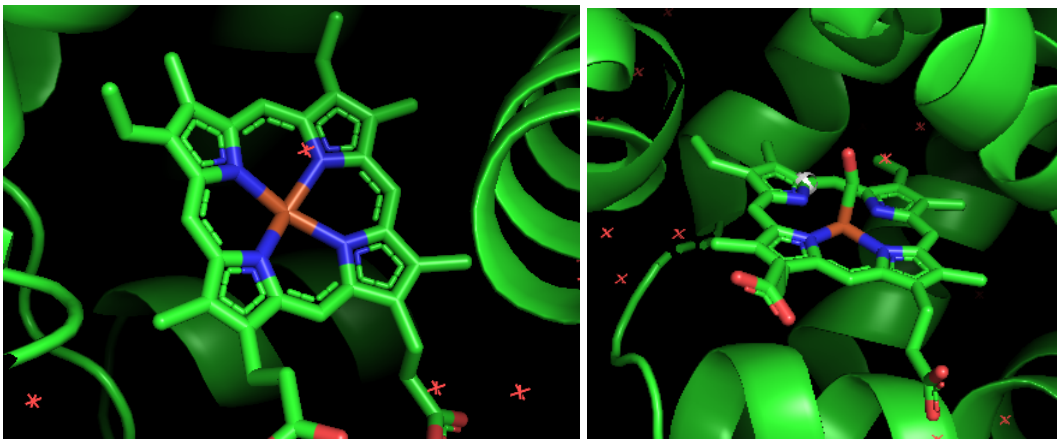
(a) (b)

Figure 1: (a) RBC with sickle (left) and normal (right) shape and (b) impact of sickle cell dieasease.

These sickle cell RBC are unable to pass through the narrow capillaries (artery and vein), resulting in painful vasoocclusive crises (Rees et al., 2010). Due to the sickle cells, hemolysis occurs which leads to anemia and a shortened lifespan of the blood cells. The only approved therapy for their remedy is the use of cytotoxic drug and hydroxyurea (Charache et al., 1995) that work via an unknown mechanism. It is also used in the case of fetal hemoglobin, which is protective against sickling of blood cells. There is the variation in the patient response to the used drugs and compliance to hydroxyurea (Charache et al., 1995). As the mutation in RBC causes sickle cell disease due to the alteration within the hereditary code of hemoglobin protein within the cell. Hence, the changed hemoglobin protein cannot carry oxygenated blood legitimately to the body. The three-dimensional structure of hemoglobin is studied at the experimental level and found clearly that the sickle hemoglobin (HbS) is unique in characteristics of being normal when it carries oxygen and abnormal when it is deoxygenated (Muirhead & Perutz, 1963; Wishner et al., 1975; M. F. Perutz & Lehmann, 1968). Various experimental works focuses on finding the close and complex relationship between this genetically infectious disease and protein dynamics (Kästner, 2011). On the other hand, heterozygotes found within the sickled gene are moderately defensive against the peril from passing on by intestinal sickness, which is presently firmly verified in therapeutic science through a number of clinical hones from distinctive parts of Africa and Asia (Luzzatto, 2012). The sickle hemoglobin protein is incapable to carry the oxygen

to the body in appropriate way, subsequently the life time of the cell is decreased. Such unusual hemoglobin protein causes the red blood cell rigid and sticky, which are unable to transport the oxygen properly (Yesudasan et al., 2019) in the blood. The disorder due to sickle hemoglobin in the human blood brings many problems in the body like painful swelling (Figure 1 (b)) of hands and feet, risk of brain stroke and increasing risk of pathogenic infections (Barpanda, 2013; Akinsheye & Klings, 2010). In fact, the mutation in hemoglobin, like the sickle cell, originated as an adaptation to protect the patient from malaria disease. This type of abnormality of hemoglobin protein normally appears in the people whose ancestors are from the former and current malaria effected zones of the world. Africa, Asia, Far East, the Middle East, Mediterranean Islands, and South America (Carter & Mendis, 2002; Weatherall, 2008)are the fertile zone for the occurrence of this diseases.

Hemoglobin is responsible in transporting oxygen, carbon dioxide, carbon monoxide and water throughout the human body. The hemoglobin carry inherited genetic character. In such hemoglobin, the sequence of amino acid residue are changed due to the incorrect form of DNA code. As a result, the flexible bi-concave structure of RBC changes to a strong elongated shape. Such a blood cell carrying less oxygen, appears fibrous and looks like a sickle shape cell. Sickle hemoglobin protein (HbS) means the damages in the sickle erythrocytes in which there is increase in the density of the damaged cell. Due to which, reduction in deformability, shortens the life span and increase in the adhesivity, (Stradner et al., 2007) of blood cell. The polymerization is the primary reason for making the deoxygenated hemoglobin (HbS) protein in to sickle cell. Figure 2 (a) shows the sixth glutamate residue of the α chain in hemoglobin protein which is undergoing a mutation into valine (Figure 2 (b)). This mutation brings the charge on the surface of hemoglobin. The sixth valine residue of the β chain in the deoxygenated HbS lies on the surface of the protein. This presence of hydrophobic residue of valine in the β chain makes a hydrophobic interaction with the neighbor β chain of other hemoglobin molecule (Stradner et al., 2007; Mosher, 1992) and so on. Figure 2 (a) and (b) represent the normal glutamate residue and mutated valine residue which reduces the solubility of the sickled protein in aqueous state.



(a) (b)

Figure 2: Heme ligands of (a) sickle hemoglobin protein with valine and (b) normal hemoglobin with glutamic acid in 6^{th} position of beta chain.

This study explores the thermal and transport properties of sickle and ordinary hemoglobin proteins. Previous studies have examined some of the mechanical, thermal, and transport properties of these proteins around the human body temperature (310 K) (Rugh & Bharathan, 2005). Computational simulations of sickle hemoglobin proteins reveal significant anisotropy and variations in strength among the four types of hemoglobin proteins. The study also shows that glycated hemoglobin proteins are generally more mechanically stiff than deoxyhemoglobin, oxyhemoglobin, and carboxyhemoglobin molecules (Yesudasan et al., 2018). The structure of hemoglobin exhibits a soft outer shell and a stiff core. The mechanism of protein misfolding in sickle cell anemia is not yet fully understood, making it challenging to develop effective anti-sickle drugs. The aggregation events of proteins that lead to tissue selective deposition in human pathology are also under investigation. Further research is necessary to fully comprehend the dynamics of protein aggregation and the structural features of the hemoglobin protein aggregate, which is crucial for a comprehensive understanding of sickle hemoglobin. Highlighting of the dynamics of mechanistic details of protein aggregation and the structural features of the aggregate of hemoglobin protein is critical for a comprehensive understanding of the sickle hemoglobin (Xue et al., 2019).

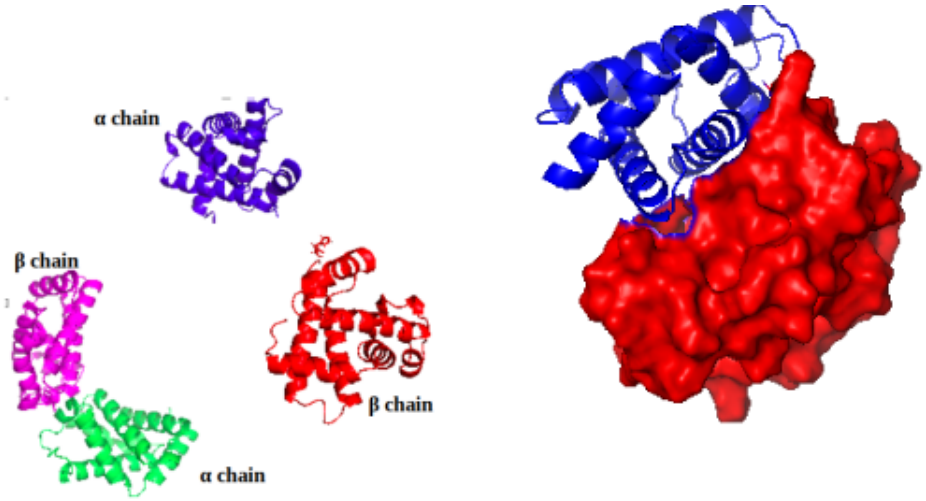
According to N.S. Akbar (2016), the use of magnetic fields and metallic nanoparticles is crucial for controlling drug delivery through narrow arteries (N. S. Akbar, 2016). The design of this control mechanism is influenced by surface tension on the artery's skin (N. S. Akbar & Butt, 2016), as explained by N.S. Akbar and Butt (2016). The impact of surface tension of blood becomes significant when there are metachronal waves created by a large number of particles on the artery walls, or when the walls undergo peristaltic motion during breathing, as described by N.S. Akbar (Abdollahzadeh Jamalabadi & Keikha, 2016; Abdollahzadeh Jamalabadi et al., 2015). In some cases, the presence

of a magnetic field can affect bio-fluid flow (N. S. Akbar et al., 2016; N. Akbar et al., 2016), such as in the processing of physiological body fluids and physical practices, where bio-magnetic fluid dynamics in stenosed arteries play a role. It has been observed that mechanical flow control devices can disrupt fluid flow patterns, leading to stent breakage or embolism in the artery's concerned path (Mann et al., 1938; Chakravarty & Mandal, 2000) demonstrated the effects of blood flow on decreasing lumen area of blood vessels, which highlights the relationship between fluid flow patterns in blood vessels and cardiovascular diseases like arteriosclerosis (Abdollahzadeh Jamalabadi et al., 2016).

1.1.1 Dimer Structure of Sickle and Normal Hemoglobin Proteins

A dimer hemoglobin molecule has two globin chains, each of them posses their own heme protein attached on them. One globular chain is alpha and the other is beta (Barrick et al., 1997). Two of such dimer hemoglobin individually behaves as functional group and forms the functional tetramer of hemoglobin protein.

There are two dimer structures of $\alpha\beta$ in both hemoglobin protein of sickle and normal (Figure 3 (a)). A dimer of sickle hemoglobin is the crystal structure of carbonmonoxy sickle hemoglobin (Figure 3 (b)). It is a biochemical function that evolves in the biological process. Two distinct polypeptide molecules (Figure 3 (b)) of chain α and chain β are present in sickle dimer (as in normal dimer) with genetic name of HBA1 and HBA2 respectively.

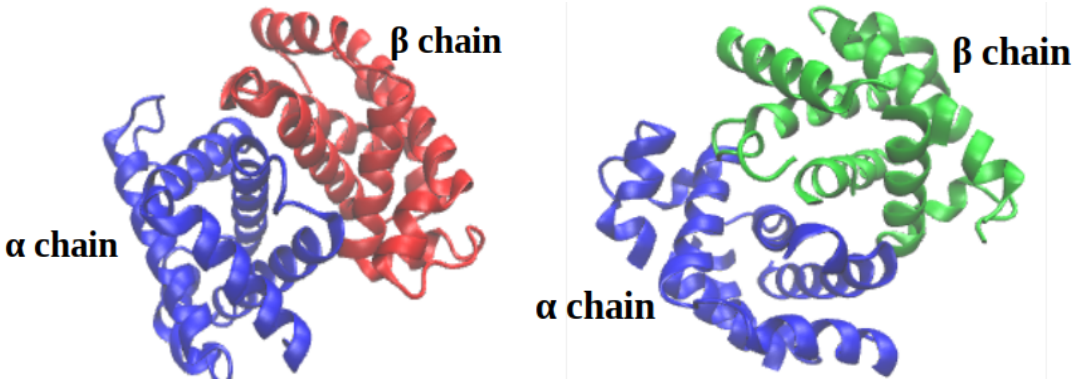


(a) (b)

Figure 3: (a) Binding site of dimer in HbS and (b) surface interaction in $\alpha - \beta$ chain dimer of sickle hemoglobin.

They have atomic weight as in normal hemoglobin protein dimer. Each chain surrounds a heme ligand with a bonding potential. The crystal structure of carbon monoxi dimer

sickle hemoglobin in R-state conformation behaves as biochemical functional group which evolves in biological process and commonly responses with hydrogen peroxidein extra cellular regions of human body (Eaton et al., 2007). Two of the dimer hemoglobin molecules combine together to form a functional tetramer. The dimers are interacts with ionic and hydrogen bonds and prevents the movement of the poly-peptide chains. It is a component of hetero tetramer structure in hemoglobin. Two distinct poly-peptide molecules (Figure 4 (a)) are bonded in this sickle dimer with genetic name of HBA1 and HBA2 respectively. The normal hemoglobin dimer (Figure (4) (b)) also consists of 141 amino acids in chain α chain and 146 residues in β chain. Disease like sickle cell anemia is poorly understood whether it is protein miss folding or aggregation during the period of last fifty years of its study Luzzatto (2012). Elucidation of the mechanistic details of protein aggregation and the, structural features of the aggregates is critical for a comprehensive understanding of the mechanisms of protein (Xue et al., 2019).



(a) (b)

Figure 4: Dimer structure of (a) sickle and (b) normal hemoglobin proteins.

The transportation of heat through the material body driven by a temperature gradient and is explained by thermal conductivity (Xue et al., 2019).

The hemoglobin comprises four sub-units, each having one poly-peptide chain and one heme protein (Figure 3 a)). All hemoglobin carry the same prosthetic heme atom, press protoporphyrin related with a poly-peptide chain of 141 (alpha) and 146 (beta) amino corrosive buildups. The ferrous particle of the heme particle is connected to the N of the histidine buildup. The porphyrin ring is infusing into its take by a phenylalanine of its poly-peptide chain. The poly-peptide chains of grown-up hemoglobin are of two sorts, one as alpha and the other beta chains, comparative in length but contrasting in amino corrosive grouping. The alpha chain of all human hemoglobin, embryonic and grown-up, are the same. The non-alpha chains incorporate the beta chain of ordinary grown-up hemoglobin ($\alpha 2\beta 2$), the gamma chain (Marengo-Rowe, 2006) of fetal hemoglobin ($\alpha 2\beta 2$), and the delta chain of HbA2. In some variants, the gamma genes are duplicated, giving rise to two kinds of gamma chains (Marengo-Rowe, 2006). The structure of

hemoglobin has been studied in depth by x-ray analysis (Borgstahl, 1992).

According to Xue (2019), thermal conductivity explains the transportation of heat within a material body, which is driven by a temperature gradient (Xue et al., 2019).

Hemoglobin is composed of four sub-units, with each sub-unit containing one poly-peptide chain and one heme group (Figure 3 a)). All hemoglobin molecules have the same prosthetic heme atom, which is connected to a poly-peptide chain consisting of 141 (alpha) and 146 (beta) amino acid residues. The ferrous particle of the heme group is attached to the N of the histidine residue, and the porphyrin ring is embedded within its take by a phenylalanine residue of its poly-peptide chain. The grown-up hemoglobin consists of two types of poly-peptide chains, namely alpha and beta chains, which are similar in length but have different amino acid sequences. The alpha chain of all human hemoglobin, both embryonic and grown-up, is the same. The non-alpha chains include the beta chain of normal adult hemoglobin ($\alpha 2\beta 2$), the gamma chain of fetal hemoglobin ($\alpha 2\beta 2$), and the delta chain of HbA2 (Marengo-Rowe, 2006). In certain variants, the gamma genes are duplicated, resulting in two types of gamma chains (Marengo-Rowe, 2006). The structure of hemoglobin has been extensively studied using x-ray analysis (Borgstahl, 1992).

1.1.2 Tetramer Structure of Sickle and Normal Hemoglobin Protein

Hemoglobin has complicated molecular structure with four protein chains liganded with four small non-protein molecules, called heme, which transports oxygen from the lungs to the rest of the body.

A tetramer hemoglobin molecule has two different physical forms, one when carries oxygen and the other when the oxygen is dissociated from it. The mutation originated polymerization in the hemoglobin distorts the red blood cells into an abnormal sickle shape (Pauling, 1964). The four chains in the hemoglobin protein consist of two alpha and two beta chains in a double dimer structure as in Figure 5(a) of alpha-beta chains. A heterozygotes structure is the mixture of normal hemoglobin A (HbA) and mutant hemoglobin S (HbS). The hemoglobin A is responsible to stops polymerization, preventing serious sickling (Murray et al., 1990). The only difference in between sickle and normal hemoglobin is the 6th residue in their beta chain (Figures 5 (b) and (c)). i.e., one of the beta chain contains valine and the other beta chain of the tetramer contains glutamic acid in their 6th residue.

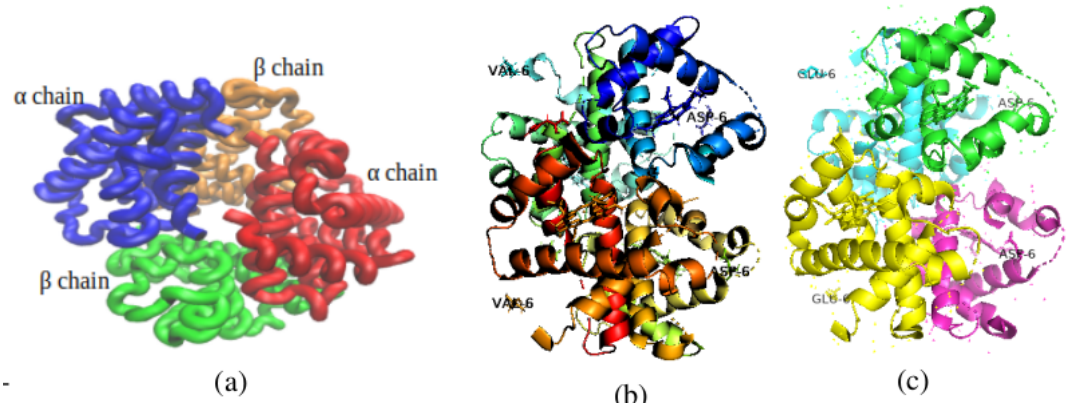


Figure 5: Hemoglobin protein with (a) double α-β dimer structure, (b) sickle hemoglobin and (c) normal hemoglobin protein.

The change in hemoglobin occurs due to the mutation of glutamic acid residue in to valine, which then generates abnormal beta globin with deformed sickle cell (Platt et al., 2000), i.e. an amino acid, called valine acid (Figure 6 (a)), is mutated into another amino acid, glutamic acid (Figure 6 (b)). This mutated tetramer Hb protein (HbS) are responsible for the red blood cells to become rigid and sticky, so that the hemoglobin protein cannot transport the oxygen properly throughout the body. It causes

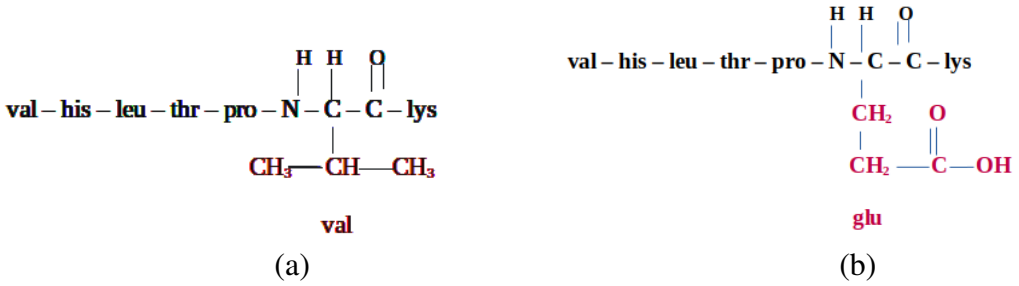


Figure 6: First seven residues with (a) valine at 6th position of beta chain in sickle and (b) glutamic acid residue at 6th position of beta chain in normal hemoglobin.

several types of disorders in the human body. The bio-physical phenomena of sickle hemoglobin protein are not studied completely yet. The further understanding of the phenomena of polymerization and the techniques for their remedies are still awaited in medical physics. Thus, our study focused on relating the structural features and the occurrence of bio-physical process using molecular dynamics.

1.1.3 Importance of Steered Molecular Dynamics (SMD)

SMD has been a powerful tool and is sensitive to the secondary structure of proteins. It has been widely used to investigate the protein unfolding pathways and their elasticity. The application of the SMD simulation in recent studies of conformational dynamics and structural characteristics of proteins has been widely used. In specific, the hydrogen

bond interaction of back bone and side chain and consolidation of the unnatural 6^{th} valine within the fibrus amino acid residue with beta chain, are the capacity of SMD to investigate the specific basic and energetic data (Xiao et al., 2019) of the protein.

The applications of these methods in recent studies of protein polymerization have also been reviewed. The constant pulling velocity and spring constant of SMD simulations performed by nano scale molecular dynamics (NAMD) is used in which external force is applied to alpha carbon atoms in the chain during N-C terminal SMD. This steered molecular dynamics allows to study the biological processes in respect of time scales. The instantaneous unbinding of ligands and conformational changes in biomolecules can be studied during SMD simulation. In which one backbone carbon atom of the alpha chain of one end is fixed and the other alpha carbon of the same chain are pulled on to unfold the protein. This dummy, as well as the SMD atoms, are moved at constant velocity and after that the interactive between both is measured utilizing the classical equation (Phillips et al., 2005). The virtual spring is considered in between the fixed atom and the SMD atom with a spring for the appropriate simulation run (Phillips et al., 2005).

1.1.4 Binding Free Energy Difference

The binding free energy of dimer molecules takes important role in the formation of molecular structure of the proteins. The poly-peptide chains of protein are joints together to form a long protein chain with a number of alpha helices and beta sheets. The binding energy of molecule is also govern by the response of the solvent to the molecule in which it immersed. The ligands are another components in hemoglobin protein which measures the binding energy of two molecules in the complexes. The binding free energy of two molecule and their contact area can report the tensile strength of the interface.

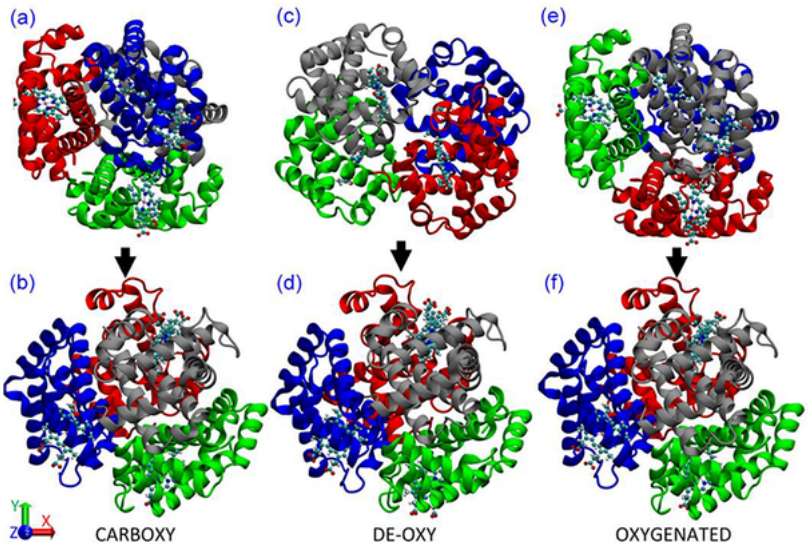


Figure 7: Tetramere view of HbS molecule (Powrel et al., 2022).

We are interested to the crystal structure of carbonmonoxy-ligated sickle Hb (COHb S) (Figures 7 (a) and 7 (b)) and the normal one (COHb A) (Figures 7 (e) and (f)), in the quarter-nary classical R-state at 1.76 Å. The structural analysis of COHb S, COHb A and Deoxy-HbS (Figures 7 (c) and 7 (d)) provides atomic level insight into the protein structure so that, why is the liganded hemoglobin does not form a polymer (Ghatge et al., 2016a). The mutation occuring in the beta chain of hemoglobin molecule changes the bio-physical properties of the entire RBC (Thom et al., 2013). The fundamentally sickle cell are in deoxygenated (T-state) structure of hemoglobin protein which distorts the cell into concavo-convex shape (Safo & Abraham, 2005; Ghatge et al., 2016b). In such case the mutation in sixth valine in the beta chain is responsible for polymerization of the sickle cell by interacting with residues residing in the adjacent hydrophobic pocket.

To understand the binding affinity of chains in dimer, we have investigated the free energy of chain α and chain β by using umbrella sampling molecular dynamics (USMD).

1.1.5 Thermal and Transport Properties

The specific heat capacity and thermal conductivity of hemoglobin protein gives the thermal property and diffusion coefficients provides the transport property of any protein. Computer simulation by NAMD can estimates the specific heat, thermal conductivity, diffusion coefficient of the protein etc. These parameter plays important role in the protein dynamic.

Specific Heat Capacity

The specific heat capacity of blood at constant volume is the measure of heat needed to raise the temperature of blood by 1^0. It is essential to understand the thermal stress and heat balance in the tissue of human body separately. A hemoglobin is combination of many elements in colloidal form. The human body is full of different kinds of tissues having different specific heat capacities. A group of workers estimated the average specific heat of human body ranging from 2.44 to 3.39 kJ kg^{-1} $^0C^{-1}$ (Xu et al., 2022).

Diffusion

Thermal diffusion occurs due to temperature differences in the phase space and go on diffusion in all directions continuously until equilibrium occurs. Therefore, the equation of continuity is the basic concept used to treat the problems of diffusion. As the heat diffuses once it never comes back, it is considered an irreversible process. The human arms and feet are consist of bone, muscle, fat, blood and skin. There are additional vital tissues exist in the lungs, abdomen and brains in the head. This model able to calculates the conduction of heat based on the temperature gradients between the tissue nodes.

As the heat flux gauge can not measures heat transfer between the surface zones and the internal body cavity experimentally. So, it is measured externally with considerable errors by computational simulation.

The blood transportation incorporates the process of mass, linear momentum, angular momentum, energy, and charge movement from one part of the body to another portion due to non-equilibrium or in-homogeneity within the system equilibrium like temperature, charge and thickness variety within the bulk framework. The physical quantity movement in the solution can be quantitatively measured by estimating the coefficient of viscosity of the protein molecules in their solution. A detail study of the diffusion of mass and heat conduction through hemoglobin protein in water is essential to understand the life processes and the physical mechanism of organic compounds. Many researches have already been focused in finding the mechanical properties like surface tension of amino acid residues constituting the hemoglobin protein (Nosal-Wiercińska, 2012; Korotchkina et al., 2004). To the best of our knowledge, the diffusion coefficient and thermal conduction of the sickle and normal hemoglobin molecules in water using MD simulation has not yet been compared yet.

Therefore, we intensify our study to these properties of hemoglobin proteins up to atomic level. Carsten Funck et al., have pointed the independent of the apparent diffusion coefficient on the diffusion time. They also observed the large differences between individual samples, with results ranging from 1.26 to 1.66 μm^2/ms for flow-compensated and 0.94 to 1.52 μm^2/ms for mono-polar gradients. Particles start to diffuse due to the density differences between the medium and go on diffusing in all directions continuously until equilibrium occurs. Diffusion is driven by a gradient in Gibbs free energy. Therefore, the equation of continuity is the basic concept used to treat the problems of diffusion. The ordinary transport phenomena illustrates the diffusion of mass, thermal transport, energy transport, and charge movement from one place of the system to the another due to non-uniformity or in-homogeneity of the system. Spread, is such an critical transport property, in which mass is transported as a result of self-assertive atomic movement (Jin & Chen, 2000). Different exploratory methods of peak-height strategy, nuclear magnetic reverberation (NMR), and MD simulation (Virk et al., 2015) are commonly used to study the diffusion of amino acids in water. This observation concerned with the effects of temperature on the diffusion of Hb (Koirala et al., 2020). As it were little, non-polar molecules,like oxygen and carbon dioxide, can diffuse effectively over the layer (Arnone, 1974). In any case, little polar atoms are transported by proteins. Subsequently, the estimation of diffusivity of hemoglobin protein is critical in planning the drugs. The arrangement of the sub units are known as the tetramer structure and differs in the oxy and deoxyhemoglobin. The de-oxy hemoglobin is stabilized relative to oxyhemoglobin (Arnone, 1974). In human hemoglobin, the binding between the

poly-peptide chains is critical because of the gap between the two poly-peptide chains in the hemoglobin molecule. It becomes narrower when oxygen molecules become attached to the ferrous atoms. This has been investigated by Max Perutz to a molecular form of paradoxical breathing. Unlike the lungs, the hemoglobin molecule contracts when oxygen enters and expands when oxygen leaves the heme. Compounds other than oxygen, such as nitric oxide and carbon monoxide, are also able to combine with the ferrous atoms of hemoglobin. Carbon monoxide are firmly attached to the ferrous atom than oxygen does. Once the carboxyhemoglobin is formed inside blood, oxygen cannot displace carbon monoxide to any extent. This forms the molecular basis of coal gas poisoning.

Interactions are responsible to make the shape and structure of the hemoglobin proteins. They also provides the solvation accessible surface area to the protein in contact with water. The contact area in between the protein chains and protein-water interface guides the folding of proteins. Similarly, the diffusivity of sickle and normal hemoglobin protein helps to understand the supply of the oxygen to blood, The specific heat capacity and thermal conduction of protein are responsible to maintain the human body temperature.

To incorporate the total elastic properties of normal and sickle hemoglobin protein, it is necessary for estimation of the forces required in separating the alpha chain from the other three chains. So, the study was focused on estimation of the force needed in breaking the hydrogen bonds of alpha chain in both sickle and normal hemoglobin protein. After serious review of information we basically focuses on non bonded interactions for the conformation of structure of sickle and normal hemoglobin protein. The stable structure of hemoglobin proteins of sickle and normal are compared for looking in to the level of interactions and active residues involved in them. Binding free energy, elastic stiffness, thermal and transport properties of both hemoglobin proteins are estimated to conclude the possible nature of anti-sickle drugs.

1.2 Rational of the Study

The hemoglobin protein with sickle cell causes the anemia (SCA) in which the red blood cell becomes sticky, stiff and non spherical shape. It is a genetically transmitted disease. When both mother and father carry one copy of sickle cell gene then their child will be affected by sickle cell disease. Though the death rate is low the patient may not active in work. Some of the places of Africa and Asia are found as infected area. Nearly 1250 patients are still present in western zone of Nepal (Pande et al., 2019).

Hundreds of years already spent in accumulating the depth knowledge of sickle cell disease. Still there are some of the unresolved problem on designing the anti-sickle drugs. Change in the heat capacity, thermal conduction, diffusivity and the binding

mechanism of alpha and beta chain in the dimer and tetramer structure of hemoglobin protein causes the sickle Hb protein. It seems essential to study the causes of these alternation in the physical parameters of sickle from normal hemoglobin and their remedy. The estimation of the free energy difference between two proteins in the hemoglobin pays great attention to understanding their binding mechanism with targets; and also the polymerization of hemoglobin as sickling. The knowledge of the transport phenomena of hemoglobin infers more understanding about transport of oxygen in the R-state and T-state of the hemoglobin protein. In this context, the study of mechanical, thermal and transport properties of mutated hemoglobin and normal hemoglobin protein are relevant to get more insight at molecular level.

The interactions makes the poly-peptide globular structure forming the alpha helix and a beta sheet of protein. Thus it looks essential to study them for all the four systems. Binding free energy estimation by umbrella sampling is the major way to understand the binding mechanism of intra-chain alpha-beta dimers in both sickle and normal hemoglobin. SMD study provides the forces required in breaking the hydrogen bonding in each of the chains in both of the tetramer structure of sickle and normal hemoglobin proteins.

1.3 Objectives of the Study

This research work is targeted at realizing the sickle hemoglobin protein by molecular dynamics simulation. The development of concepts of physics in molecular dynamics is highly used to understand the active mechanisms of proteins in different branches of the natural sciences. However, in sickle cell hemoglobin protein we have not seen sufficient work up to the molecular level. Therefore, we have limited the study under the following general and specific objectives.

General Objective

To study the mechanical, thermal and transport property of mutated hemoglobin protein of sickle cell.

Specific Objectives

To fulfill the general objectives following specific objectives are made,

- To study about the elastic properties of sickle and normal hemoglobin proteins.
- To estimate the binding free energy and RMSD value of sickle and normal hemoglobin proteins.
- To estimate the specific heat capacity of sickle and normal hemoglobin proteins.

- To estimate the transport properties like coefficient of diffusivity of sickle and normal hemoglobin proteins.

1.4 Organization of the Thesis

The structure of this thesis is organized as follows:

- In Chapter 1, we presented the motivational part of this research work which provides the understanding of sickle hemoglobin protein in structural details.
- In Chapter 2, we present literature which provides a background of the research gap to set up the objectives of this research work.
- In Chapter 3, "Materials and Methods" describes the theoretical background required to complete the research work. It includes in brief the SMD of the beta chain of Hb to explain the elastic phenomena of the beta chain, classical molecular dynamics relation to estimate the free energy, Fick's law to estimate the self diffusion coefficient, Darken's relation to estimate the binary diffusion coefficient and the effect of temperature on diffusion coefficient. We mainly discussed the binding free energy and umbrella sampling methods to estimate the free energy of the dimer of sickle and normal hemoglobin protein.
- The main findings of this work are presented and discussed in Chapter 4, entitled 'Results and Discussion'. Section 4.1 introduces the background for the whole chapter. We discuss the structural stability of proteins and analyze the calculation of RMSD of different protein molecules, the estimation of diffusion coefficients with their temperature dependency, binding free energy and elastic properties by estimating the force required in breaking the H-bonds in both the alpha and beta chain in hemoglobin proteins.
- We briefly summarize the conclusions about possible application of the work in Chapter 5. Finally, the summary is presented in chapter 6, followed by the references.

CHAPTER 2

2. LITERATURE REVIEW

2.1 Genetral Consideration

In this chapter, we briefly discuss the literature relevant to our works. First we review the literature related to sickle cell hemoglobin proteins followed by their structural formation, difference in between sickle hemoglobin and normal hemoglobin proteins. How is the sickle cell hemoglobin detected at first and then their chronological order. The tetramere and dimer structure of hemoglobin protein and the study of their sequences. We then present the literature which focused on the mechanical, thermal and transport properties of sickle and normal hemoglobin. We also review the free energy difference between alpha and the beta chain in hemoglobin protein structure. We also discuss the literature related to the effect of temperature on diffusion coefficient as well as molecular dynamics study. Finally diffussion property and diffussion coefficient of both normal and sickle hemoglobin protein are discussed for looking into the transport property.

A mutation in normal hemoglobin occurs by changing one of the bases in the DNA sequence from the adenine to thaymin. This alter the amino acid residue of the haemoglobin protein from glutamic acid to valine. Moreover, there are many findings that confirm the differences in sickle hemoglobin in comparison to normal are due to ionizable amino acids (Wells & Itano, 1951; Eaton, 2002). The gene causes the human body to generate a new type of haemoglobin called sickle hemoglobin (HbS) and behaves very differently to regular mechanism of haemoglobin (HbA) (Guo et al., 2010). Mutated sickle hemoglobin govern the red blood cells to develop abnormally and inforce to sickle shaped instead of the usual doughnut shape. Sickle cell mutations have different hemoglobin which is polymerizes in long fibers, as the red blood cell loses their oxygen supply (Pawliuk et al., 2001). This polymerisation is the major factor that explore the transformation of RBCs into sickle-shaped deformed floppy discs of protein. They are harder and less flexible and can stick in the blood vessels during transportation, causing blockages. During the last few decades, a lot of research work has been done on hemoglobin proteins for their physiochemical properties. Sickle cell disease was detected

at first by James B. Herrick in 1910. It was nearly 4 decade later, Linus Pauling and his colleagues concluded that the sickle cell disease was caused by a genetic disorder in the blood protein. It was one of the un-known molecular disorder (J. B. Herrick, 1910) and was not sure at that time whether the blood condition was a diseased or a manifestation of another disease (J. Herrick, 1924) in the patient body. Fibrinogenation of protein was studied at first in 1943 by Bailey, Astbury & Rudall (Bailey, 1948). Pauling, L. et al studied the molecular disease of sickle cell anemia in 1949 (Pauling et al., 1949). These serial studies were reviewed by Linus Pauling, who was the first to hypothesize in 1945 that the disease might originated from the abnormality in the hemoglobin molecule (Pauling, 1964). The polymerization of hemoglobin occurs in three stages:viz nucleation, fiber growth, and fiber alignment. A double nucleation mechanism phenomenon is described for the polymerization of sickle hemoglobin. This mechanism postulates two pathways for polymerization. Research study indicating that, the polymerization is initiated by homogeneous nucleation in the solution phase of protein formation and then followed by nucleation of additional polymerisation on the surface of an existing polymer called heterogeneous nucleation (Ferrone et al., 1985). As the surface area of polymers is increased continuously with time, the heterogeneous nucleation provides a mechanism for the extreme autocatalysis and is manifested as an apparent delay in the kinetic progress curves. In this process, each spherulitic domain of already formed polymer is considered to be initiated by a single homogeneous nucleation event and then heterogeneous. The end stage of polymerisation is a complicated structure of helical fiber (Van Eps & De Jong, 1997). The increased stiffness of such helical HbS fibers are considered to be the main reason for the wide variety of shape of the protein that deoxygenated RBC's (red blood corpuscular) from the SCD patients (Ferrone & Rotter, 2004; Gulbis et al., 2005; Maciaszek & Lykotrafitis, 2011) . The polymerized HbS fibers are stiff, they create protrusions and causes vesiculation of the RBC membrane (Statius van Eps, 1999) (Li & Lykotrafitis, 2011). Proteins have the primary structure of a linear chain of amino acids. The secondary structure of hemoglobin is comprises with the stabilized hydrogen bonds between the atoms in the polypeptide backbone. Tertiary structure is the three-dimensional shape of the protein chains determined by the stabilized interactions between the side chains in the molecules. The quaternary structure of protein is the association between two or more polypeptides chains. But every protein may not have a quaternary structure (Kappel et al., 2019; Banach et al., 2019). Delay-time of gelation of the three mutants of r HbS were studied in 1.8 M potassium phosphate (pH 7.34) with a temperature jump from 4^0 C to 30 ^{0} C (Briehl, 1982). The increase in delay time over that of r HbS was observed, which also confirm an overall decrease in the polymerization of the three mutants of HbS. A more detail and quantitative investigation has also been carried out to determine the equilibrium solubility in 0.1 M potassium phosphate (pH 7.35) at 25^0 C of the three HbS mutants as well as of mixtures of these mutants

with HbS versus mixtures of fetal hemoglobin (HbF) and Hb A with HbS (Ho et al., 1996). The inhibition of polymerization demonstrated that the interactions in between the two amino acid residues, alpha 114Pro and beta 87Thr, are very important to the formation of HbS polymer. Modification of these amino acid results in an anti-sickling potential (Ho et al., 1996) in the protein molecules. Previous studies have demonstrated that the unique behavior of hemoglobin S with respect to polymerization, molecular orientation (Ho et al., 1996), and decreased solubility consequent upon deoxygenation is responsible for the sickling phenomena. The blood from the patient with sickle cell disease shows the increased number of sickle cells, increased erythrocyte mechanical fragility and increased whole blood viscosity at the reduced oxygen situation (Ho et al., 1996). Furthermore, the rate of gelation and sickling are found to be sensitive to very small change in the physiologic parameters (Allison et al., 1956; Eaton et al., 1976; Harris & Bensusan, 1975). The erythrocytes of sickle-cell anemia contains the hemoglobin S in sufficient concentration to allow the changes to occur within the physiologic range of oxygen tensions. They also contain a concentration of hemoglobin S, insufficient to show these changes except upon extreme reduction in the oxygen intake, to levels usually considered nonphysiologic. When the fibrus RBC of blood is exposed with normal oxygen environment, (normoxic) they melts the fibers. But the repeated exposure exhibits the permanent and irreversible damage during the low-oxygen and normoxicenvironments (Evans et al., 1984; Itoh et al., 1995; Schmalzer et al., 1987; Alexy et al., 2006) . This may lead to the consequence of abnormal entry of Ca into the sickle red blood cell (Maciaszek & Lykotrafitis, 2011). The occurence of the irreversible membrane damage causes an increase in the stiffness of the sickle cells, and becomes unable to unsickle after re-exposure to room atmospheric air (irreversibly sickled cells) (Kaul & Xue, 1991; Kaul et al., 1983). These RBCs may adversely affects the blood flow (Carr & Cockelet, 1981) and contributes to the vessel occlusion, poor oxygen transport, and hemolysis (Kato et al., 2017; Merkel et al., 2011). Vasoocclusion is the major cause of morbidity and mortality in sickle cell disease. As hydroxyurea therapy technique may hold promise for assessing changes in blood viscosity in sickle cell and other hematological diseases, other techniques are also awaited for drugs (Hosseinzadeh et al., 2018). Although various techniques can be used for the measurement of rheological properties of biological fluids like human hemoglobin, acoustic levitation is an unique technique which provides rheological information for biological samples in a minimally invasive way only by using one single drop of the sample (Hosseinzadeh et al., 2018). On the other hand, the heterozygotes of the sickle genes are relatively stronger in fighting against the malaria, as now firmly established through a number of clinical field studies from different parts of Africa and Asia including Terai region of Nepal) (Shrestha & Karki, 2013). The thermal and transport property of human body and its constituent tissue is of new scope of study. The latest study estimated out the average specific heat capacity

of human body is to be approximately 2.98 kJ kg^{-1} $^0C^{-1}$ (Xu et al., 2022). However,the study of the specific heat capacity, thermal conductivity and difussion of sickle cell hemoglobinis protein is still lacking in molecular level.

Proposed study is of the estimation of free energy, diffusion coefficient and elastic properties of both hemoglobin proteins using molecular dynamics simulations. These studies are helpful in the applied field of drug discovery, which is a newly generated scope so far. In this context, molecular dynamics study may play a crucial role in the determination of the diffusion coefficient of both Hb proteins in water, and hence, to improve the pharmaceutical scope.

Molecular Dynamics

Molecular dynamics (MD) simulations basically predict the motion of the atoms in protein molecules on a time scale using the basics of Newton's law by considering the interatomic interactions (Karplus & McCammon, 2002). The simulation can snapshot a wide variety of important biomolecular processes, including conformational change, ligand binding, and protein folding and unfolding, revealing the positions of all the atoms at femtosecond level of resolution. In 1957, Alder and Wainwright at first studied the lysine brush and therapeutic semax peptides based on molecular dynamics simulations (B. J. Alder & Wainwright, 1957; Ercolessi, 1997). They studied phase behavior of hard sphere systems from simulation carried out in a rectangular box under periodic boundary conditions. Gibson and co-workers used a finite difference method to solve equations of motion during simulations (Gibson et al., 1960). Rahman, in 1994, reported the diffusion coefficient of liquid argon estimated from the molecular dynamics simulations performed with pairwise Lennard Jone potential (Rahman, 1964). They carried out the simulation in a cubical box of size 10.229 σ under the periodic boundary conditions in which they estimated the value of the diffusion coefficient of liquid argon at 94.4 K is 2.43×10^{-6} cm^2/sec using mean square displacement (MSD) of particles. In this work, Rahman also explained the pair correlation and auto correlation functions between argon atoms. For hard sphere, the detail about the velocity autocorrelation function was explained by Alder and Wainwright (B. Alder & Wainwright, 1967). The dynamical behavior of pancreatic trypsin protein was studied under molecular dynamics simulation by (McCammon et al., 1977). Similarly, Levitt and Warshel performed simulations to understand the protein folding mechanism (Levitt & Warshel, 1975).

Research Gap

After reviewing the relevant literature, it has been observed that many experimental works were carried out to understand the solubility and diffusivity of hemoglobin protein at body temperature. The crystal structure of hemoglobin has already been

reported (Eaton et al., 2007). Similarly, optimized geometry as well as the different spectrum of hemoglobin were studied by different groups using DFT and NMR calculations. To our knowledge, although experimental studies have been going on about the calculation of diffusivity, thermal conduction and other thermodynamic parameters, the binding free energy and mechanical properties as well as the transport properties of hemoglobin in aqueous medium have not been performed at molecular level. The estimation of the free energy difference between two proteins in the hemoglobin pays great attention to understanding their binding mechanism with targets; and also the polymerisation of hemoglobin as sickling. The knowledge of the transport phenomena of hemoglobin infers more understanding about transport of oxygen in the R-state and T-state of the hemoglobin protein. In this context, the study of mechanical, thermal and transport properties of mutated hemoglobin and normal hemoglobin protein are relevant to get more insight about the solvation process of human hemoglobin at molecular level. The disease of sickle cell is studied almost for 100 years and confirm as genetic disease. The recommended treatment for sickle cell disease is the gene therapy (Orkin & Bauer, 2019).

CHAPTER 3

3. MATERIALS AND METHODS

3.1 General Consideration

In this chapter, we have discussed the theory and methodology used to accomplish the objectives of our research work. This section consists of three subsections, viz., the first part (subsection 3.2) is for theoretical background in which we discuss the free energy estimation by statistical method, specific heat capacity, thermal conduction and diffusion of protein molecules in water. We have also discussed the application of Fick's law of diffusion, Einstein's equation to measure the diffusion coefficients and the effects of system temperature on diffusion coefficient. Similarly, the mid part (subsection 3.3) includes the theory of molecular dynamics in detail in which we have discussed bonded and non bonded interactions, the pattern of bond formation and the nature of the bonds and non-bonded Coloumb and Lennard-Jone's interaction potentials and the theory of SMD. The third part (subsection 3.4) comprised computational details for system setup and simulation. In this subsection, we have addressed the relation of theory with the simulation of protein molecules. However, the systems of sickle and normal hemoglobin protein for required calculations designed for simulation purposes are elaborated on in this section. At last, we have discussed the analytical tools of root mean square deviation (RMSD), hydrogen bonding, salt bridge, hydrophobic, van der Waals (vdW) and electrostatic interactions, solvent accessible surface area (SASA) and umbrella sampling (US).

3.2 Theoretical Background

In this section, we have briefly described the theory related to free energy, specific heat capacity and diffusion. The theoretical section also includes the theory of molecular dynamics (MD), SMD and umbrella sampling.

3.2.1 Free Energy

The free energy of any system is the amount of energy available to do work by the system. Helmholtz (H) and Gibbs (G) free energy can be defined in terms of internal energy (U) and entropy (S) as,

$$\text{A(N,V, T)= U-TS.} \tag{3.1}$$

$$\text{And G(P,V,T)=A+PV} \tag{3.2}$$

Generally we can measure the free energy difference between two thermodynamics states. The difference in Gibb's free energy between two thermodynamic state is,

$$\triangle G = \triangle A + P \triangle V \tag{3.3}$$

In this equation, $\triangle A$ is the change in Helmholtz free energy and $\triangle V$ is the change in volume. Energy is always associated with binding the molecules in a complex. An amount of energy is released when the molecules are dissociated, as free energy. The binding free energy is the fundamental property of the multi molecular system (Wong & McCammon, 2003). It defines the free energy difference between the bound and completely unbound states of the dimer or tetramer molecules. The bound state is the fully consistent structure whereas, the completely unbound state is any of the intermediate states in which the interaction between the molecule is negligible. The free energy difference of dissociation of the molecule is essential for structure-based drug design.

The numerical solution method is commonly used for free energy estimation. Among the various method, the non-Boltzmann sampling method is also used for the calculation of the free energy of molecules (M. Allen & Tildesley, 1987) computationally. The free energy can also be estimated theoretically by considering the thermal and conformational contributions (Powrel, 2022). Although, the free energy is commonly estimated by decoupling method, the method of umbrella sampling is considered for hemoglobin like larger protein (Kästner, 2011). The pathway method is another way for finding the mechanistic and kinetic information of the binding system. This method measures the potential of mean force (PMF). It is defined as the negative logarithm of the probability at a certain value of the reaction coordinate. Gibbs's free energy is expressed as (Gilson et al., 1997).

$$\Delta G = -k_B\, T\ \ln \frac{Q(N,V,T)}{Q_0(N,V,T)} \tag{3.4}$$

where Q and Q_0 are the partition function of two different states. Then, these samplings

of each window are executed for the MD simulation under the application of a harmonic potential between the two states of a sample under the applications of appropriate force constant (Kästner, 2011) of biasing. The free energy difference of the systems of $\alpha\beta$ dimer in hemoglobin is obtained by the relation of (Murcko, 1995; Raha & Merz Jr, 2005)

$$\triangle G = \triangle G^{\alpha} + \triangle G^{\beta} - \triangle G^{0} \tag{3.5}$$

where $\triangle G^{0}$ is the free energy of the $\alpha\beta$ dimer complex, $\triangle G^{\alpha}$ is the free energy of the alpha chain and $\triangle G^{\beta}$ is the free energy of the beta chain in the dimer complex. The systems of sickle and normal hemoglobin dimer of $\alpha\beta$ complexes are set in the VMD at first for making the required translations. Every successive window was prepared, translating the center of mass (COM) distance of β chain by 1 Å along the negative x-direction by taking the COM distance of α chain as the reference. The window size of each system confirms the required overlapping of the successive windows to incorporate all the reaction coordinates of space. The reaction coordinates of umbrella sampling are the distance between the center of mass (COM) of α and β chain. If the system is lying along the direction of the negative x-axis, then to make the required overlapping of reaction coordinates, a bias potential of V (ξ) is applied to force the system for fluctuating in the coordinate space along the given direction, which is given by the Equation 3.4 (Kästner, 2011).

3.2.2 Specific Heat Capacity

The specific heat capacity is another thermodynamic property that the material takes response in the structural stability during the heat transfer. Thus in molecular mechanics, the specific heat capacity at constant volume is (Phillips et al., 2005),

$$c_v = \frac{\langle E_B^2 \rangle - \langle E_B \rangle^2}{k_B T^2} \tag{3.6}$$

where, $\langle E_B^2 \rangle$ is the average of the square of the total energies and $\langle E_B \rangle^2$ the square of the average total energy and k_{B} the Boltzmann constant.

3.2.3 Diffusion

Diffusion is the process of transfer of matter from the region of higher concentration of particle to lower concentration and is defined by using Fick's law (Crank, 1979) as,

$$J = -D_{self} \nabla\, C(r,t) \tag{3.7}$$

Here, C(r,t) is the concentration of the particles as a function of position and time and D_{self} is the diffusion coefficient. The negative sign in Equation (3.5) indicates that the diffusion increased along the direction opposite to that of the increased concentration. These two functions, flux and concentration can be measured in terms of the same quantity of displacement. D_{self} has the dimension of (length)^2 (time^{-1}). This measures the 24rate of diffusion of the substance. The self-diffusion coefficient is generally estimated by using Einstein's relation. According to this relation, the macroscopic parameter, diffusion coefficient (D_{self}) for a three-dimensional system is calculated from the microscopic quantity of the mean square displacement (MSD) (M. P. Allen & Tildesley, 2017a),

$$D = Lt_{t\to\infty}\left(\frac{\langle|r(t)|^2\rangle}{6t}\right) \tag{3.8}$$

Moreover, the binary diffusion coefficient of substance in a mixed solution is again obtained from Darken's equation of diffusivity (Darken et al., 2010).The generalization of the Darken relation is,

$$D_{ij} = [x_i/(x_i + x_j)]D_{j,self} + [x_j/(x_i + x_j)]D_{i,self} \tag{3.9}$$

where, x_i and x_j qre the mole fraction of species i and j respectively (Krishna & Van Baten, 2005).

3.3 Mpolecular Dynamics

In this work, we have carried out molecular dynamics (MD) study of sickle and normal hemoglobin proteins to understand their structural, mechanical, thermal and transport properties at the molecular level. Molecular dynamics is a statistical approach in which different properties of our interest can be estimated from the analysis of the trajectories followed by the particles in 6N-dimensional phase space. It is a deterministic technique in which we solve the equation of motion to generate the trajectory. During simulation, we first assign a position for each particle and the new position can be generated by solving the equation of motion (M. P. Allen & Tildesley, 2017b) as,

$$m_i\frac{\partial^2 r_i(t)}{\partial t^2} = F_i\{r_i(t)\} - \gamma_i\frac{dr_i(t)}{dt}m_i + R_i(t) \tag{3.10}$$

The term $m_i\frac{\partial^2 r_i(t)}{\partial t^2}$ represents the net force experienced by an atom in the molecule, $F_i\{r_i(t)\}$ is the total force experienced by the atom due to the neighboring atoms. $\gamma_i\frac{dr_i(t)}{dt}m_i$ is the damping force term with damping constant γ_i m_i and the final term $R_i(t)$ represents the random force acts on the atom.

Also, the force can be expressed as a negative gradient of potential (V_i) given by,

$$F_i = -\frac{dV_i}{dr_i} \tag{3.11}$$

In equation 3.11, the potential V_i, is the sum of bonded and non-bonded potentials involved in bonding the system, i.e.,s,

$$V_{total} = [V_{bonded}] + [V_{nonbonded}]$$

Bonded Interactions

The bonded potential is the sum of bond stretching, bond angle as well as the contribution of potential due to dihedral interaction (Phillips et al., 2005). All the above mentioned potentials are harmonic except for proper dihedral which has periodic variation. So, the bonded potential can be expressed as,

$$V_{bonded} = [V_{bond} + V_{angle} + V_{improper} + V_{proper}].$$

The bond stretching potential between two particles i and j separated at a distance r_{ij} is defined as,

$$V_{bond}(r_{ij}) = \frac{1}{2}k^b_{ij}(r_{ij} - b_{ij})^2$$

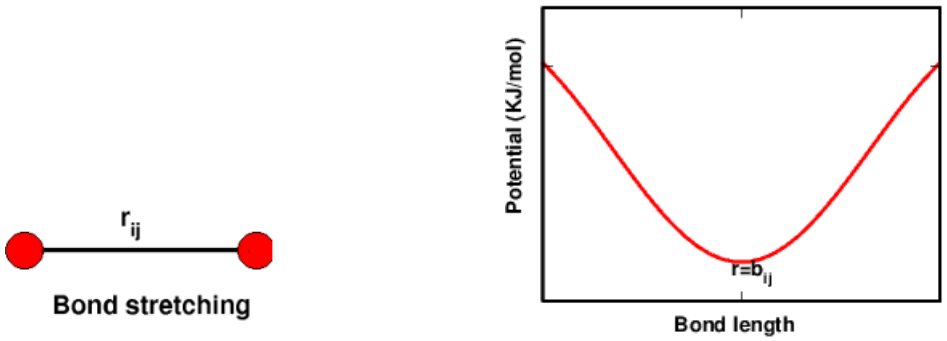
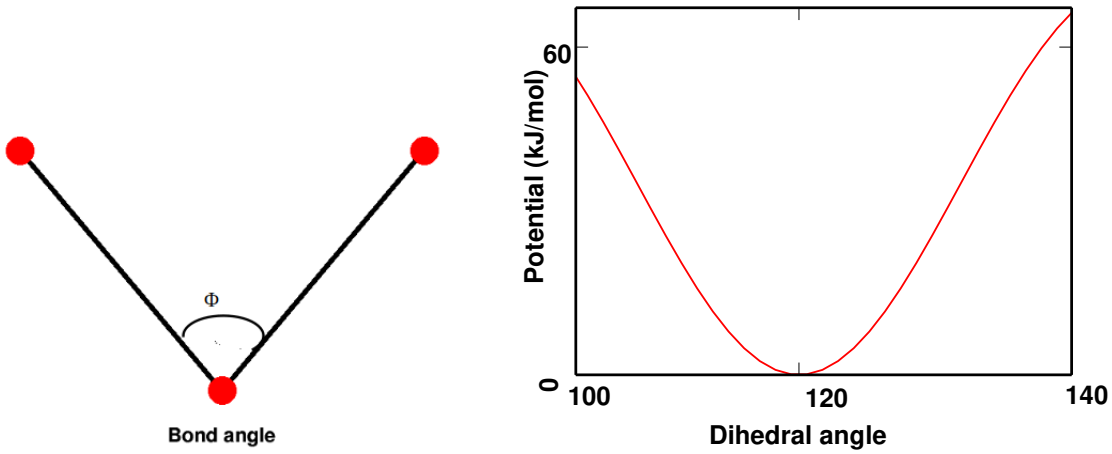


Figure 8: (a) Bond stretching structure and (b) bond stretching potential.

where, b_{ij} is the equilibrium bond length and k^b_{ij} is the force constant. The variation of harmonic potential of bond stretching is shown in figure 8 (b). Again the bond angle potential is expressed as,

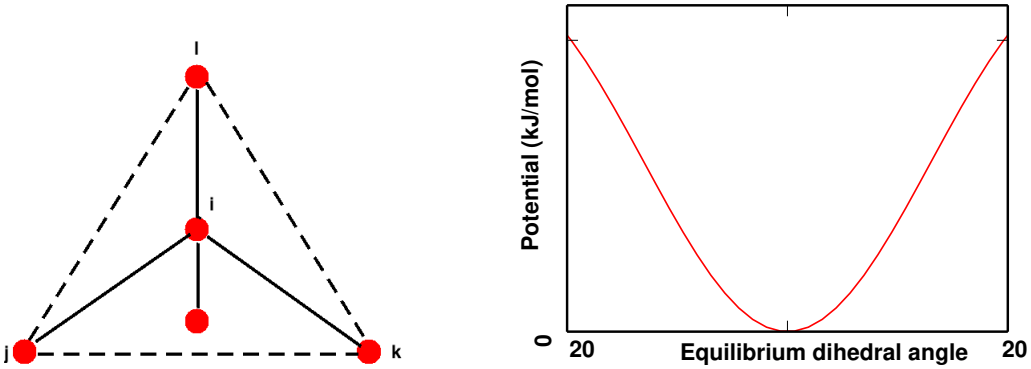
$$V_{angle}(\theta_{ij}) = \frac{1}{2}k^{\theta}_{ijk}(\theta_{ijk} - \theta^0_{ijk})^2.$$



(a) (b)

Figure 9: (a) Bond angle structure and (b)bond angle potential.

Here, θ_{ijk} is the equilibrium bond angle and k^{θ}_{ijk} is the force constant. The bond angle θ_{ijk} is oscillating harmonically around equilibrium bond angle of θ^{0}_{ijk} . The variation of the bond angle potential is as shown in figure 9 (b). The improper dihedral potential energy is,



(a) (b)

Figure 10: (a) Improper dihedral structure and (b) improper dihedral potential.

$$V_{improper}(\xi_{ijkl}) = \frac{1}{2}k_{\xi}(\xi_{ijkl} - \xi_0)^2.$$

Here ξ_0 is the equilibrium dihedral angle and k_{ξ} is the constant. The improper dihedral potential is harmonic in nature as in Figure 10 (b). The periodic potential of proper dihedral is defined as,

$$V_d(\Phi_{ijkl}) = k_{\Phi}[1 + cos(n\Phi - \Phi_s)], \tag{3.12}$$

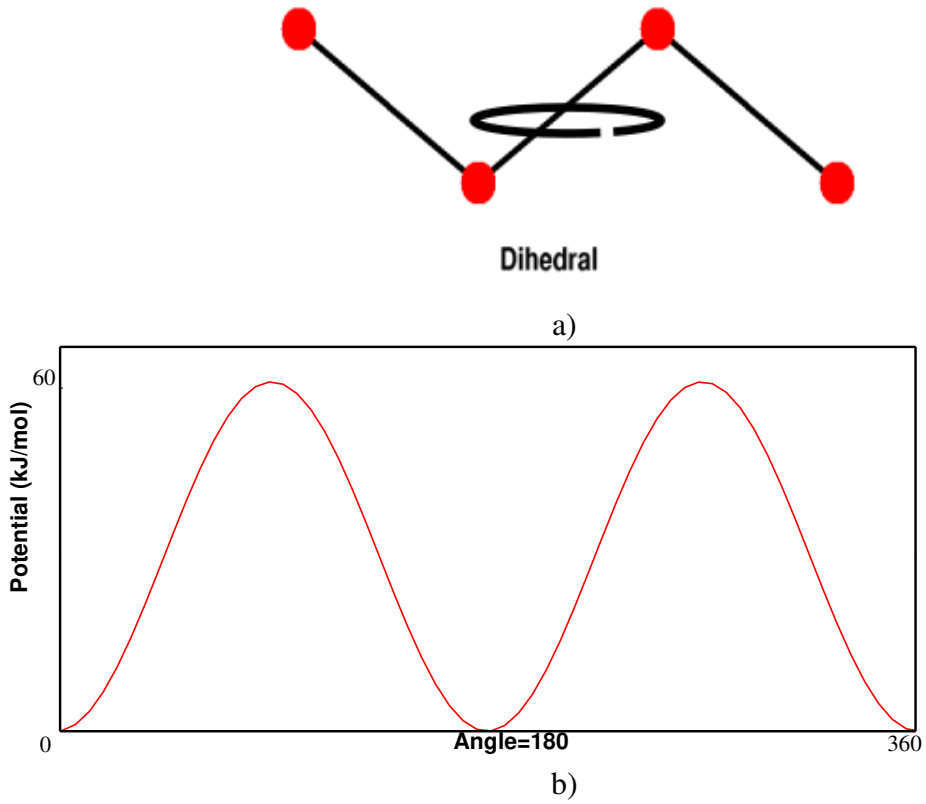


Figure 11: a) Proper dihedral structure and b) periodic dihedral potential.

here, Φ the angle in between the planes ijk and jkl and k_Φ the constant. The proper dihedral (Figure 11 (a)) potential is periodic in nature as in Figure 11 (b).

Non-bonded Interactions Electrostatic and van der Waals forces are essential for both intramolecular and intermolecular binding in biomolecules. Hemoglobin relies on these interactions to bind its beta and alpha chains when forming its quaternary structure (Ghatge et al., 2016b). Some amino acids such as aspartic acid and glutamic acid carry a negative charge, while others like lysine, arginine, and histidine have a positive charge (Fischer et al., 2002). Additionally, atoms within the molecule carry partial charges, which can form electric dipole moments and contribute to particle binding through electrostatic interactions (González-Ruiz & Gohlke, 2006). While the Lennard-Jones potential may not be effective in containing systems with strong localized bonds or delocalized electron seas, it remains useful for understanding fundamental concepts in condensed matter physics through ion retention.

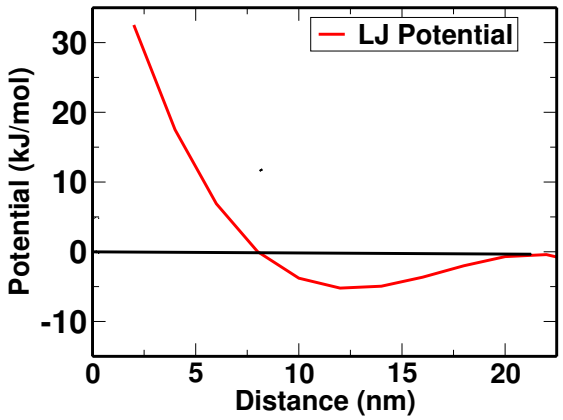


Figure 12: Lennard-Jones potential

The L-J potential for interaction (Figure (12)) between a pair of atom is given by (Lennard-Jones, 1931),

$$V_{LJ} = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right], \tag{3.13}$$

where ε represents the depth of the potential well, σ refers to the inter-particle separation representing the strength of the interaction where the potential changes the sign of positive and negative and defines the length scale. The positive sign of the term $\frac{1}{r^{12}}$ indicates repulsive and dominates at a short distance it is due to the non-bonded overlap of electronic clouds explained by Pauli's exclusion principle. Unlike, the term $\frac{1}{r^{6}}$ dominates at the long distance and the negative sign indicates the attractive potential due to van der-Waals dispersion forces caused by the dipole-dipole interaction.

The another non-bonded interaction arises due to some partial charges in the atoms or molecules is Coulomb electrostatic interaction. The interaction between two point charge q_i and q_j is defined by Coulomb potential as

$$V_{coloumb} = \frac{q_i q_j}{4\pi\epsilon_m r} \tag{3.14}$$

Total interaction energy due to electrostatic and van der Waals interaction is given by (Lennard-Jones, 1931),

$$V_{non-bonded} = [V_{LJ} + V_{coloumb}] \tag{3.15}$$

$$Or, V_{non-bonded} = \frac{q_i q_j}{4\pi\epsilon_m r} + 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \tag{3.16}$$

where, ε_m denotes the permittivity of the medium and q_i and q_j are the two charges

seperated by distance r_{ij}. Thus the total potential energy in the molecule is considered as,

$$V = \frac{1}{2}k_{ij}^{b}(r_{ij} - b_{ij})^2 + \frac{1}{2}k_{ijk}^{\theta}(\theta_{ijk} - \theta_{ijk}^{0})^2 + \frac{1}{2}k_{\xi}(\xi_{ijkl} - \xi_0)^2 + k_{\Phi}[1 + cos(n\Phi - \Phi_s)]$$

$$+4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_m r_{ij}} \tag{3.17}$$

The potential energy function consists of four terms that account for different types of oscillations and interactions between atoms. The first term represents bond oscillations around the equilibrium bond length with a specific bond force constant. The second term describes bond angle oscillations around the equilibrium angle with a specified angle force constant. The third and fourth terms account for improper dihedral potential and proper dihedral potential, respectively, with factors such as amplitude, periodicity, and phase. The interatomic interactions are approximated using the 12-6 Lennard-Jones potential, representing attractive and repulsive forces between atoms. The Coulombic potential describes the electrostatic interactions between pairs of atoms as point charges. Hydrogen bonding is also considered through the Lennard-Jones and Coulombic potentials. To conduct molecular dynamics simulation, these parameters are computed using quantum-mechanical calculations and compared against experimental data.

Periodic Boundary Condition

The periodic boundary conditions are inserted so as to repeat the boxes of small sizes in Figure 13 (a). The switch distance and their cut-off values accumulate the limitations of the distances to be measured in the calculations (Figure 13 (b)). The term pair list distance limits the movements of the atoms in the system. A cubical three dimensional box of specific size comprising with water and hemoglobin protein molecule is shown in Figure 13 (c).

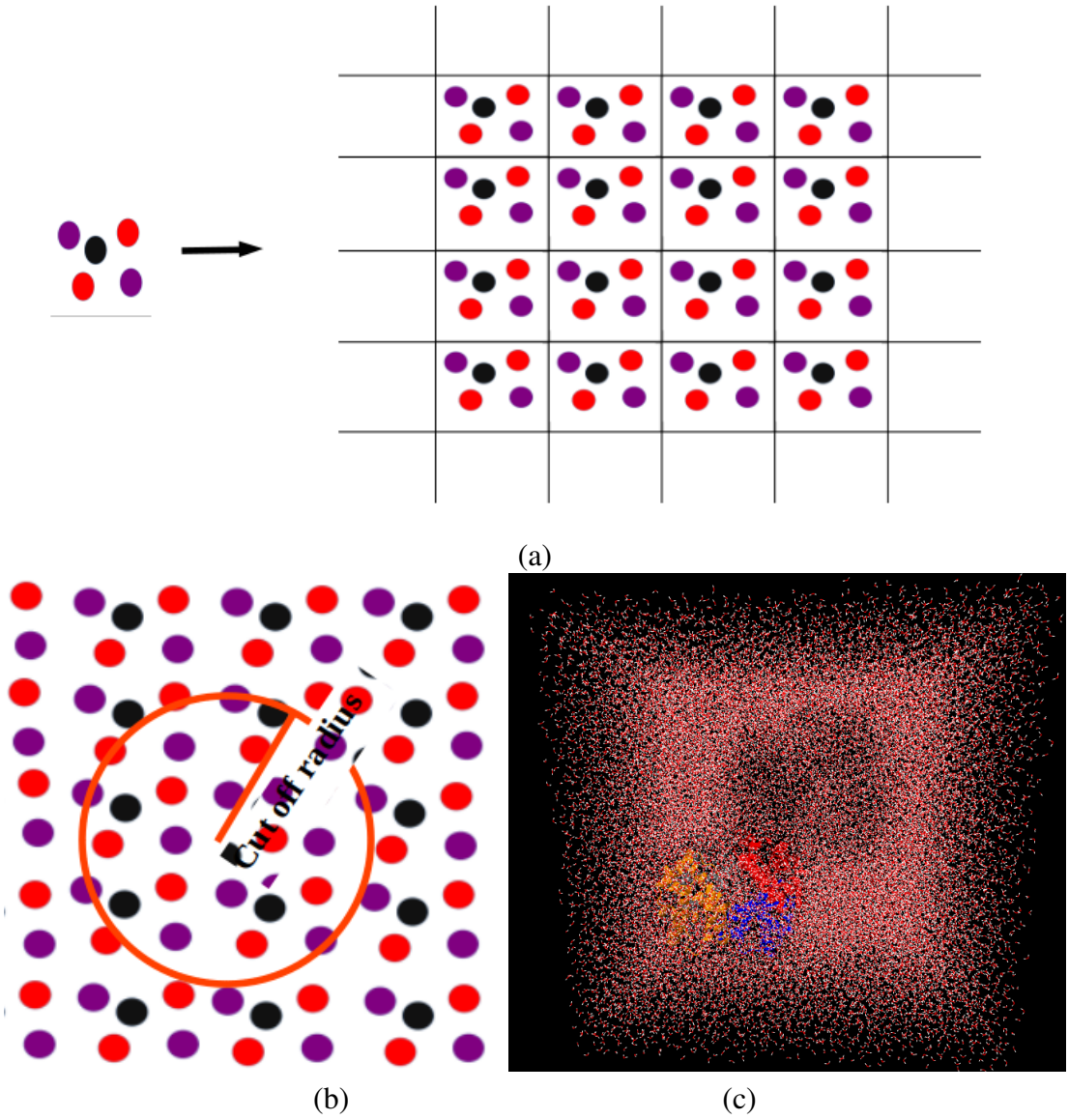


Figure 13: (a) Representation of two-dimensional simulation boxes with molecules in a repeated way applying periodic boundary conditions and (b) consideration of cut-off for interaction and (c) Simulation box with protein molecule and TIP3P water.

Initialization

NAMD configuration file (.conf) generally incorporates all the initialization parameters in the form of a ".tcl" script. A typical ".conf" file holds information about molecular structure (.psf) and coordinates (.pdb) files. A target temperature is given for maintaining the temperature of the bath. Sequentially, the initial velocity is provided for the atoms. The restart frequency and time steps provided give the trajectory path. Force field parameters and topology files carry the information of bonding in the system of molecules to exist.

Force Calculation by Integration

Force calculation is the major action of simulations in which suitable periodic boundary condition, switching of the distance and cut off distances are implemented. To generate 31new position and velocity for each particle, we solve equation of motion. Several algorithms are available to solve the equations of motion. Here we have briefly introduced Verlet and Verlet based Leap-Frog algorithms. According to Verlet algorithm, the position of each particle after each time step can be expressed in terms of the old position using Taylor expansion as

$$r(t+\delta t) = r(t) + v(t)\delta t + [F(t)/2m]\delta t^2 + [(\delta t)^3/3!]d^3r/dt^3 + O(\sigma t)^4 \tag{3.18}$$

$$or,\ r(t-\delta t) = r(t) - v(t)\delta t + [F(t)/2m]\delta t^2 - [(\delta t)^3/3!]d^3r/dt^3 + O(\sigma t)^4 \tag{3.19}$$

Here δt is the time step in the molecular dynamics simulation, m is the mass and F is the force acting on the particle. Now, evaluating the equation of motion using Verlet algorithm as,

$$r(t+\delta t) + r(t-\delta t) = 2r(t) + [F(t)/m](\delta t)^2 + O'(\delta t)^4 \tag{3.20}$$

$$or,\ r(t+\delta t) = 2r(t) - r(t-\delta t) + [F(t)/m](\delta t)^2 + O'(\delta t)^4 \tag{3.21}$$

It shows that the estimate of the new position contains an error that is of order $(\delta t)^4$. It is also seen that the Verlet algorithm does not use the velocity to compute the new position. For finding velocity we have,

$$\begin{aligned} r(t+\delta t) - r(t-\delta t) &= 2v(t)\delta t + O''\delta t^3 \\ or,\ v(t) &= \frac{r(t+\delta t) - r(t-\delta t)}{2\delta t} - O'''\delta t^2 \end{aligned} \tag{3.22}$$

This verlet algorithm seems time reversible. However, to calculate the more accurate velocity and hence its kinetic energy another algorithm of Leap-Frog is used.

Leap-Frog Algorithm

Among the different Verlet like algorithms, one is the Leap-frog algorithm (van Gunsteren, Wilfred F and Berendsen, Herman JC, 1988). This algorithm solves velocities at half integer time steps and uses these velocities to compute the new positions of the particle. To derive Leap-frog algorithm from Verlet algorithm, let us write the velocity

at half-integer time steps as follows,

$$\mathbf{v}\left(t - \frac{\Delta t}{2}\right) = \frac{\mathbf{r}(t) - \mathbf{r}(t - \Delta t)}{\Delta t} \tag{3.23}$$

$$and\ \mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} \tag{3.24}$$

From Equations (3.23) and (3.24), we can obtain the expression for new position from old position and velocity as,

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}\left(t + \frac{\Delta t}{2}\right)\Delta t \tag{3.25}$$

On using Verlet algorithm, the expression for the velocity based on the old velocity can also be obtained. Using Taylor expansion on velocity about t we get,

$$\mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}(t) + \frac{\mathbf{F}(t)}{2m}\Delta t \tag{3.26}$$

$$and\ \mathbf{v}\left(t - \frac{\Delta t}{2}\right) = \mathbf{v}(t) - \frac{\mathbf{F}(t)}{2m}\Delta t \tag{3.27}$$

Subtracting equation (3.27) from (3.26), we get

$$\mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t - \frac{\Delta t}{2}\right) + \frac{\mathbf{F}(t)}{m}\Delta t \tag{3.28}$$

Using Equation (3.28), we can obtain the new velocity of each particle from its old velocity. Although the Leap-frog algorithm is based on Verlet algorithm, the position as well as velocities are not calculated at the same time. As a result, the kinetic and potential energy is also not obtained at the same time, and hence the total energy cannot be computed directly in the Leap-frog algorithm.

After modeling all the systems, following steps are performed as in the flow chart Figure 14.

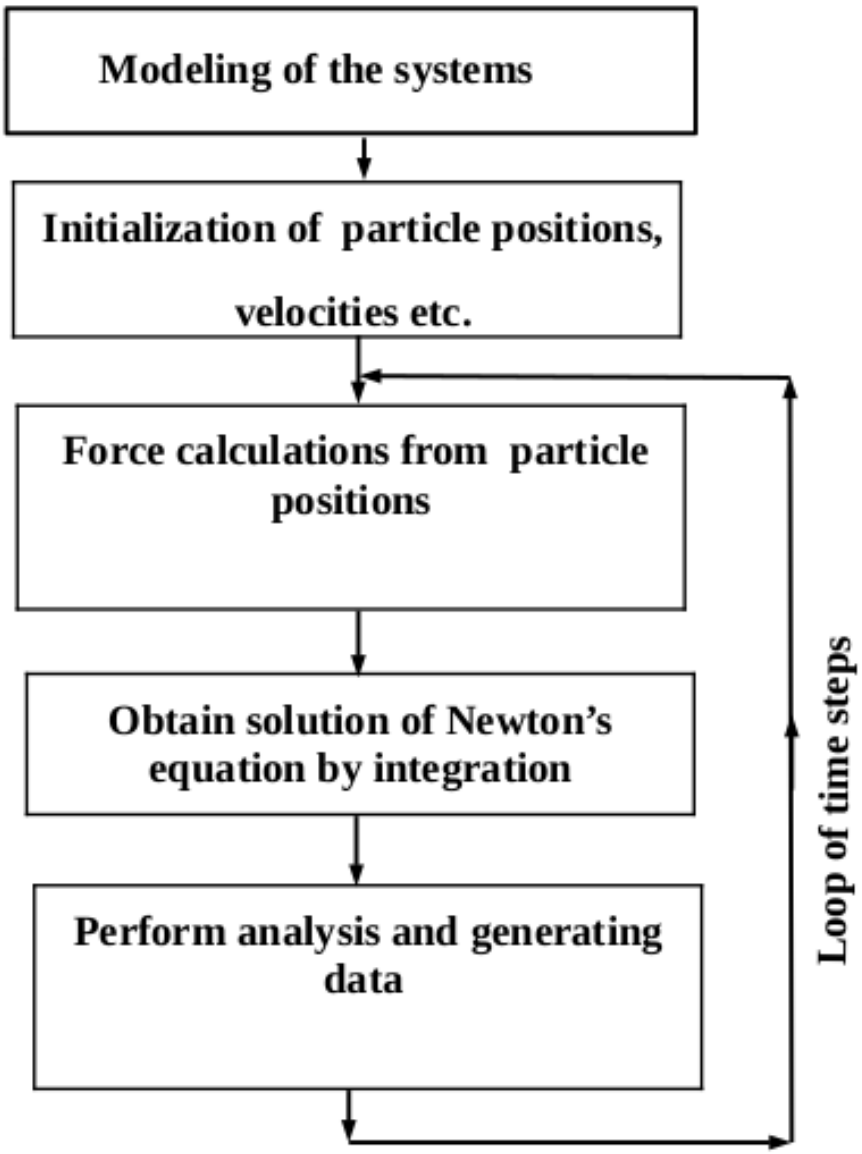


Figure 14: Flow chart of steps followed in molecular dynamics simulation.

Steered Molecular Dynamics (SMD)

Steered Molecular Dynamics (SMD) technique is a powerful tool sensitive to the secondary structure of proteins and has been widely used to investigate protein unfolding pathways. In particular, the hydrogen bonds of the backbone and side chain of protein and the non bonded interactions between the chains in the dimer protein brought the necessity of the SMD to obtain the specific structural and dynamic information (Xiao et al., 2019). The applications of these methods in recent studies of polymerization of proteins as linear or helical fibrils have also supported the use of SMD. NAMD is the freely available software that can run the constant velocity NC terminal SMD simulation. An external force is applied to one or more atoms in the protein chain, which we refer to as SMD atoms. It allows for observing the mechanism happening during SMD in time

scales like unbinding of ligands and conformational changes in biomolecules. It explore the path way of biological processes on time scale accessible in molecular dynamics simulations (Ercolessi, 1997). SMD of beta and alpha chain in the tetramer structure of normal and sickle hemoglobin protein are designed one by one. Before running the actual simulations, we prepared the system in water. During the SMD simulation one atom of the beta chain is fixed and pulled the other atom from another end to stretch and unfold the protein (Phillips et al., 2005) chain. Stretching of the beta (Figure)15) and alpha chain done through the constant velocity pulling method of N-C terminal SMD. The Figure (15) represents the SMD of beta chain in the tetramer of sickle hemoglobin protein.

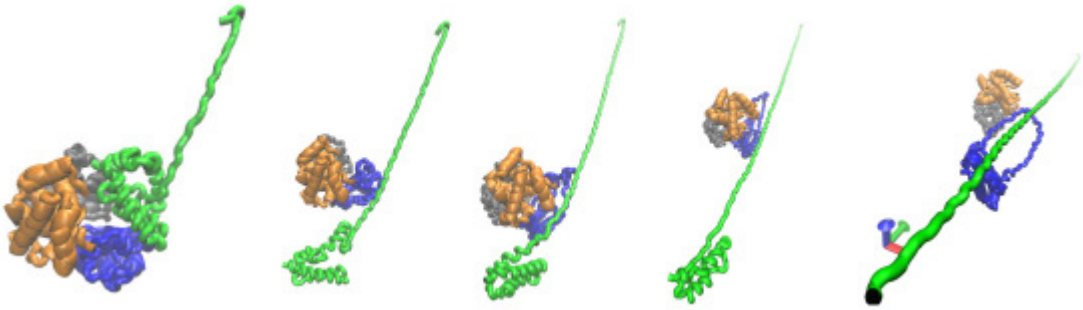


Figure 15: Snap shots of five consecutive frames of N-C terminal SMD of beta chain in sickle hemoglobin (Powrel & Adhikari, 2022).

In this type of simulation, the SMD atom is attached to a dummy atom via a virtual spring attached in it. This dummy atom is displaced at a constant velocity along a direction and then the force between the atoms is measured using the equations (M. Allen & Tildesley, 1987) of,

$$\overrightarrow{F} = -\overrightarrow{\nabla} U. \tag{3.29}$$

(Callister, 2007). For instance, unbinding of ligands and conformation changes in biomolecules. Pulling the SMD simulations at constant velocity protocol where the force experienced by the pulled terminal residue is,

$$U = \frac{1}{2} k \left(vt - (\overrightarrow{r} - \overrightarrow{r_0}) \cdot \widehat{n} \right)^2, \tag{3.30}$$

where k is the spring constant, v pulling velocity, t the time of pulling, r and r_0 are the actual position and initial position of the SMD atom. The N-C terminal SMD explores the elastic property of materials. It also explores the pathways of unfolding of protein and shows the occurrence of biological processes on time scale.

3.4 Computational Details

Computational Details For every system of considerations, separate computational details are described for obtaining the structure stability, elastic properties of alpha and beta chains, specific heat capacity, diffusivity, binding free energy and thermal conductivity of sickle and normal hemoglobin proteins. Following the computational methods we have generated their corresponding data for analysis. Two kinds of systems were generated one for tetramer study and the other for dimer study as follows:

Force Field Consideration

To describe the time evolution of bond lengths, bond angles and torsions, also the non-bonding van der Waals and elecrostatic interactions between atoms the force fields are used. A force field is a computational method used in estimating the forces between atoms within molecules and also between the molecules (Lee et al., 2016). It is also a collection of equations and associated constants designed to reproduce molecular geometry and selected properties of tested structures. To explain the time evolution of bond lengths, bond angles and torsions, also the non-bonding van der Waals and elecrostatic interactions between atoms force fields are widely used. In class one force field dynamics of particles is describe by AMBER, CHARMM, GROMOS, OPLS force field. For small molecules OPLS-All-Atom (OPLS-AA) is used wherease for larger molecules CHARMM is used.

In molecular dynamics, CHARMM force field are used for convenient as normal hemoglobin solution contains 69 132 atoms. It includes the atoms of specific distances and calculates the required parameters for analysis in molecular dynamics. The system structure of sickle and normal hemoglobin tetramer protein is obtained from the Research Collaboratory for Structural Bioinformatics (RCSB.org) (Berman et al., 2000; Brooks et al., 1983). The solved PDB files are then used in simulation by NAno scale Molecu- lar Dynamics (NAMD) (Phillips et al., 2005) and Visual Molecular Dynamics (VMD) (Humphrey et al., 1996) software. The specific heat capacity of sickle and normal hemoglobin protein are estimated. All the four systems of hemoglobin molecules are obtained from RCSB.org as follows,

System Setup of Hemoglobin Protein Tetramer

Two systems are designed by considering the hemoglobin molecules and then solved inside water in the box, cut off distance at 12.0 Å , at pressure of 1 bar and temperature 310 K using CHARMM (Chemistry at HARvard Macro molecular Mechanics) force field (CHARMM 36m).

- The sickle hemoglobin protein is obtained with the PDB ID of 2HBS.

- Normal hemoglobin protein is obtained from the PDB ID of 1A3N.

The box sizes of the MD simulations might be too small to incorporate all the atoms of protein and water for hydrophobic stabilization, so it was taken as large as possible (El Hage et al., 2018) of $(100 \times 100 \times 500)$ Å ^{3} with around 96000 water molecules and enough Na^+ and Cl^- ions to yield the system of 0.15 m/L molar concentration. Initially, MD simulations were performed in NPT ensemble with a standard requirement of at least five water molecules between any protein and the box boundary (Brooks et al., 1983). The long range electrostatics interactions are included with the consideration of Particle-Mesh Ewald (PME) with a grid spacing of 1 Å and a cutoff distance of 12 Å together with a 10 Å of switching distance for the Lennard-Jones (LJ) interactions (Stradner et al., 2007). The NAMD embedded LINCS algorithm (El Hage et al., 2018) was used for constraining the hydrogen bonds involved in hydrogen atom (Koirala et al., 2021). Each of the system is minimized at the energy level with time steps of 5000000 with each time step of 2 fs using steepest descent method. The system sample is heated by considering the Langevin Dynamics from 0 K to 310 K in the increment of 10 K in NVT for 10 ns and production with NPT, at 1 atmosphere pressure for 100 ns. The center of mass of the protein was considered as the center of mass of the simulation box. The Velocity re-scaling (El Hage et al., 2018) with 0.1 ps and Parrinello Rahman (Bussi et al., 2009) methods were used for temperature and pressure control respectively. The velocity re scaling method is an extension of the Berendsen thermostat to which a stochastic force has been chosen such as to generate a correct canonical distribution (El Hage et al., 2018). The MD simulation for all systems was carried out at 1east 1 ns at constant temperature and pressure of 310 K and 1 atm. respectively with temperature fluctuation of 0.1 K within 1 s trajectory.

Finally, the systems are further modified for NVT simulation run in which the temperature coupling feature of NAMD is implemented to ensure the temperature of the molecules.

System Setup for Dimer Hemoglobin Protein

Dimer system of hemoglobin consists of two protein chains and . The two systems of dimer hemoglobin are designed (Berman et al., 2000) as follows:

- We have chosen the R-state conformation of carbonmonoxy sickle hemoglobin dimer structure from PDB ID 5E6E and its mutated form.
- Another structure of normal hemoglobin was made from the 5E6E.pdb structure by changing (Powrel et al., 2022) the sixth residue (VAL6) of beta chain into GLU6 using CHARMM GUI.

For both of the structures, the position and coordinate (psf and pdb) were generated from the CHARMM-GUI web server (Lee et al., 2016) by inserting the topology information

from the modified CHARMM36 (i.e. CHARMM36m) force field. Then, each complex was solvated under TIP3P (Jorgensen et al., 1983) water model in the cubical box of dimensions ($90 \times 90 \times 90$) $Å^3$ (Figure 13 (c)). The periodic boundary conditions (PBC) in both the molecular systems are set in the configuration files. To neutralize the electrically charged system, we drop Na^+ and Cl^- ions into the system of 0.15 m/L molar concentration.

For the estimation of diffusivity, we have imparted the computational as follows. The already prepared dimer structure of sickle and normal hemoglobin protein systems are solved in TIP3P water for investigating diffusion. A cubical box of size 200 Å with water and protein at four different temperatures of 306 K, 308 K, 310 K, and 312 K are considered. The required simple point charge TIP3P water and modified CHARMM36m force field is used in the simulations. CHARMM36m modified embedded software provides the require bonded and non-bonded interaction parameters. The parameters for TIP3P water model is included in the INPUT file of GROMACS 5.1.1.27 (Berendsen et al., 1995; Lindahl et al., 2010; Lee et al., 2016). The Coulomb interaction is regarded as the partial charge that already exists in atoms and molecules since the same atom might have various partial charges depending on the group of attachment. The induced dipole that developed in the molecular system is where the van der Waal's interaction originates. The coordinates provided for the molecules in the pdb file might not be the equilibrium structures; rather, they are the starting requirements that are obtained from the electron probability density map created by nuclear magnetic resonance (NMR) or x-ray diffraction (XRD).

3.4.1 Simulation Details

Now, we have discussed the major steps of Molecular Dynamics Simulations. For the simulation purpose, systems are modeled by considering the suitable size of the box. A box is a three dimensional container holding specific types of water and molecules in a certain ratio.

For Tetramer Hemoglobin

The tetramer system simulation is conducted by first generating the system's force field using suitable input files in CHARMM-GUI, which includes topology and parameters. The effect of solvent layers on the stability of hemoglobin is analyzed by examining the RMSD graph obtained from the molecular dynamics simulation. The simulation is performed using NAMD package version 5.1.4 with the CHARMM modified 36m all-atom force field (Jo et al., 2008) and TIP3P water (Phillips et al., 2005) on a GPU server, as the system contains several thousand atoms. Energy minimization is conducted to eliminate undesirable fluctuations and unwanted solvent factors in the system. Particle

Mesh Ewald method is utilized to address possible long-range interactions with a non-bonded cutoff of 12 Å . Equilibration is implemented under the NVT ensemble with a temperature of 310 K and a time step of 2 fs per step. The production run is then executed for 100 ns under NPT ensemble with a time scale of 2 fs per step. The molecule-solvent ratio and system box size are maintained to ensure at least five water molecules between any protein and the box boundary (Brooks et al., 1983). Velocity is re-scaled with 0.1 ps for temperature control under Langavin Dynamics, and Parrinello Rahman method (Bussi et al., 2009) is used for pressure control. The velocity re-scaling method (El Hage et al., 2018) used is an extension of the Berendsen thermostat, with a stochastic force chosen to generate a correct canonical distribution(El Hage et al., 2018).

For Dimer Hemoglobin

In most cases, the binding free energy of dimer compounds is estimated using the umbrella sampling technique. We used this method to calculate the binding free energy of the dimer in sickle and normal hemoglobin dimer proteins in their relaxed (R-state) dimer form. Chain beta is manually moved along the negative x-direction to prepare the samples, using chain alpha as the reference. The reaction coordinate was determined to be the center of mass (COM) between these two chains. The initial sample's COM distance between the chain was calculated to be 23.8 Å . Then, for the second sample, the COM separation distance was increased by 1 Å from the first window, and so on for each successive window.

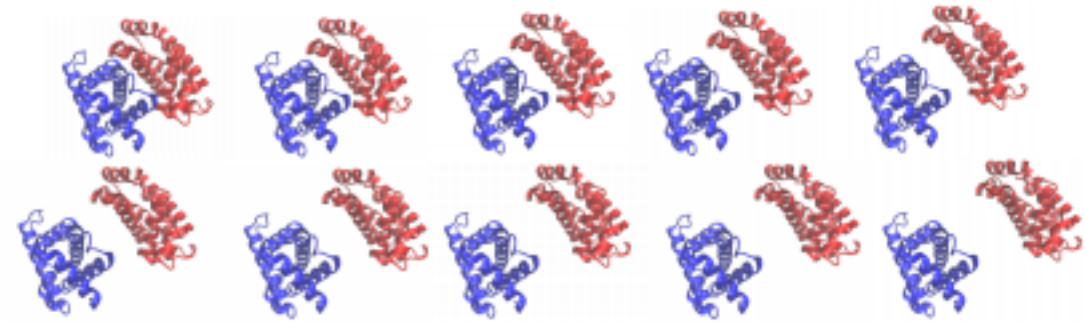


Figure 16: Ten consecutive snapshot of USSMD of dimer structure in sickle hemoglobin (Powrel et al., 2022).

For each of the two systems, twelve windows are prepared in this manner independently (Figure 16). The final window was of 12^{th} . It was chosen in such a way that there is almost no significant interactions between the α and β chain for both the systems. After suitable simulation run of 10 ns trajectories are obtained. The generated data are interpreted in WHAM analysis.

The molecules in any system are subject to random obstruction, which could cause un needed system agitation. Therefore, energy minimization is important to eliminate the impacts and keep the system at the least energy state using Steepest-descent algorithm. For each of such window, energy minimization was carried out for 50000000 steps. After minimizing the energy in the system, the dynamic properties can be studied. For protein dynamic, diffusion and viscosity usually vary with temperature, pressure and density. So, before running the simulation, the following parameters should be kept constant to make the system in a state of thermodynamic equilibrium. The equilibration process was carried out for 100 nanoseconds under NVT ensemble at 310 K and 1 atm pressure with a time scale of 2 fs per step. Then, a production run was conducted for 20 nanoseconds under NPT mode using appropriate collective variables and a time of 2 fs per step. During the simulation, the system was kept at equilibrium using the NPT ensemble with a velocity re-scaling thermostat and Berendsen barostat with specific coupling times. The isothermal compressibility used was 4.6 bar. To include Coulomb and Lennard-Jones interactions, a cutoff distance of 12 Å was used, while long-range Coulomb interactions were included using the Particle-mesh Ewald (PME) method with a Fourier spacing of 0.12 nm. The hydrogen bonds were converted into constraints using the Linear Constraint Solver (LINCS) algorithm during the equilibration process. The initial velocities of the molecules were generated using the Maxwell-Boltzmann distribution. The Leapfrog algorithm was used to calculate the positions and velocities of the atoms after given time steps. The equilibration run was performed for 50 nanoseconds with a time step of 1 fs, using an isothermal compressibility of 4.6×10^{-5} bar 1. The production run was limited to 20 ns for each window and executed for 50 ns to obtain the protein diffusion coefficients under the NVT ensemble (Powrel et al., 2022). In the NVT run, there was no need to generate initial velocities, as the velocities generated during the equilibration run were used for the Langevin dynamics.

Software Packages

The suitable software's are chosen for visualization of molecules and residues, simulation run and analysis purposes. NAMD (Phillips et al., 2005) and GROMACS (Lindahl et al., 2010) are used for molecular dynamic simulation.

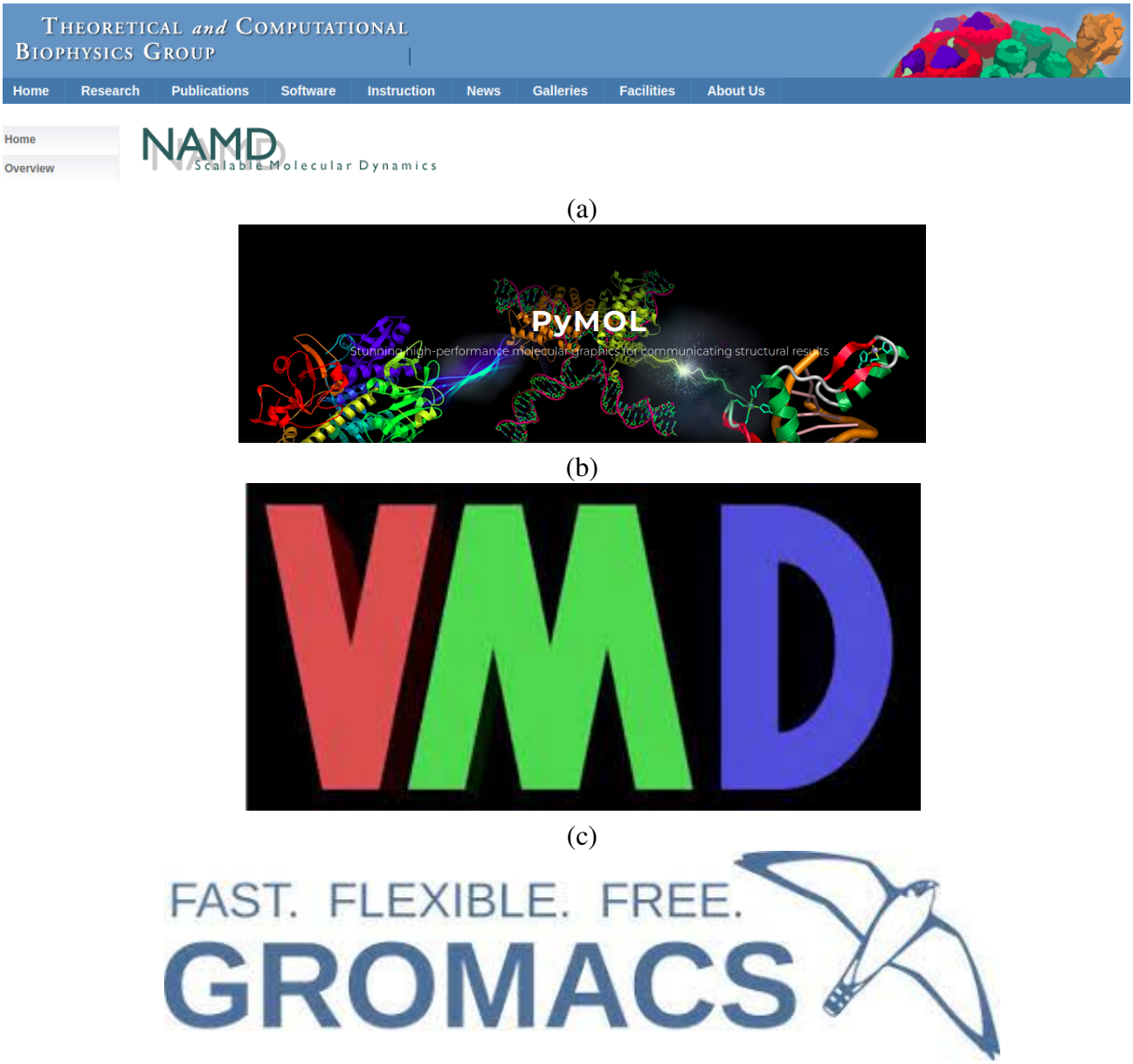


Figure 17: Software packages of (a) NAMD, (b) PyMOL (c) VMD and (d) GROMACS.

VMD (Humphrey et al., 1996) and PyMOL are used for visualization of structures. However, PIC.mbu, and WHAM (Kumar et al., 1992) are used in analysis of hydrophobic interaction and umbrella analysis respectively.

Analytical Tools

In this subsection, we briefly discuss the tools used for analysis of the trajectories obtain from the simulations.

Root Mean Square Deviation (RMSD)

RMSD estimation is the observation of the level of equilibration of the system. i.e., it gives the validity of the conformation of the system. The deviation of the molecules from the defined position in the space is essential to get the actual system. The RMSD graph carries the clue of the occurrence of the reaction in chemical bonding. Higher deviation from the mean position of the graph means the less stability of the system. RMSD values are calculated for all the atoms in the backbone of the hemoglobin protein and for the protein excluding the last five residues (Phillips et al., 2005). The numerical measure of the structural difference in between the two states given by,

$$\mathrm{RMSD}(t_j) = \sqrt{\left[\frac{\Sigma_{\alpha=1}^{N_\alpha}(\overrightarrow{r}_\alpha(t_j)-\langle\overrightarrow{r}_\alpha\rangle)^2}{N_t}\right]} \tag{3.31}$$

with,

$$\langle\overrightarrow{r}_\alpha\rangle = \frac{1}{N_t}\Sigma_{j=1}^{N_t}\langle\overrightarrow{r}_\alpha(t_j)\rangle \tag{3.32}$$

where N_α is the number of atoms whose positions are being compared, N_t is the number of time steps over which atomic positions are being compared, $\overrightarrow{r}_\alpha(t_j)$ is the position of atom α at time t_j and $\langle\overrightarrow{r}_\alpha(t_j)\rangle$ is the average value of the position of atom to which the positions $\overrightarrow{r}_\alpha(t_j)$ are being compared (Phillips et al., 2005).

Hydrogen Bonds

The number of hydrogen bonds formed throughout a simulation trajectory are counted. As the hydrogen bond formation is instantaneous so, their average frequency and occupancy percentage gives the strength of this interaction in the molecular binding. It is useful in representing the elastic stiffness of the molecules.

Salt Bridges

Salt bridges are accounted by considering the inter atomic distances between the atoms. One atom from acidic residue and the other from basic residue at minimum cut off separation are identified as salt bridges. The average occupancy of the salt bridges present are estimated to evaluate the binding strength between the molecules.

Hydrophobic Interaction

Non-polar molecules are segregated in their solution during simulation of the molecule () (Tina et al., 2007). The tendency of such aggregation in solution counts as the hydrophobic interaction in the protein dynamics which are also major driving force in biological

system. The number of residue pairs involved in hydrophobic interaction are averaged out to account the strength of this interaction in the bio-molecules.

van der Waals and Electrostatic Interaction

The van der Waals potential, on the other hand, offers energy in all kinds of atoms to give the strength to bond the molecules in the complex (DiStasio et al., 2014). The energy of the Couloumb interaction in between two charges of Q and q is determined by the Equation (3.13). The potential energy provided by the van der Waal's interaction between a pairs of atom is given by Equation (3.14).

SASA Analysis

Solvent accessible surface area (SASA) calculations is another significant metric we computed to comprehend protein folding and unfolding quantitatively. The area of a molecular surface that is sufficiently exposed to solvent molecules to interact with them is referred to as the solvent accessible surface area. When analyzing the structural stability of the molecular complex, the contact surface area between one beta globin and the other three (two alpha and one beta) is helpful. For this, the average solvent accessible surface area (SASA) of each molecule and their complex was calculated as follows:

$$\text{Contact area (S)} = \left[\frac{S_1(t)+S_2(t)-S_3(t)}{2}\right] \tag{3.33}$$

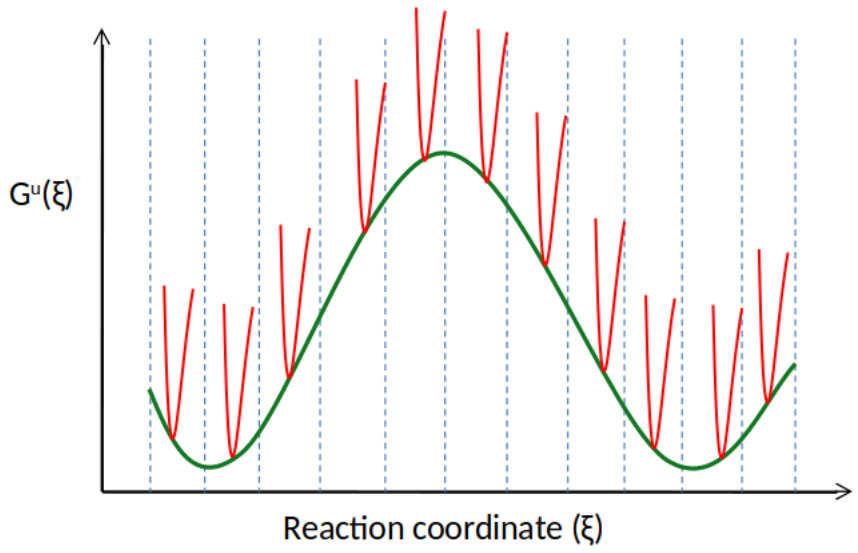
Umbrella Sampling

The umbrella sampling is a suitable techniques for estimating the free energy. We have chosen the Umbrella Sampling (US) method which is a reliable techniques for bio molecular study (Kästner, 2011). The study of binding free energy among the molecules during the unfolding of different chains can reveal their unbinding mechanism, which gives the binding strength between the molecules. The method of umbrella sampling is introduced by Torrie and Valleau in 1974-1977 for studying the bindings mechanism. In this method blue moon ensemble approach along with some restrained factor of biasing potential instead of reaction coordinate is used. The biasing potential is the harmonic one called umbrella potential as,

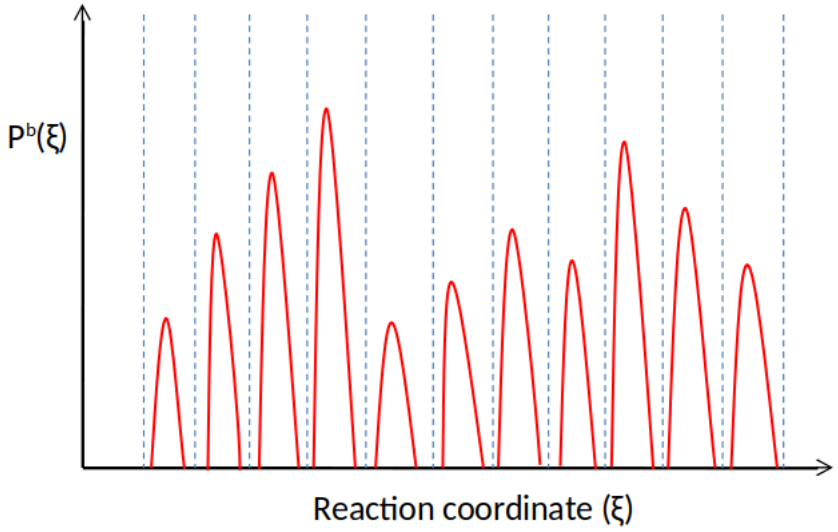
$$V(\xi) = \frac{1}{2}k(\xi - \xi_i)^2 \tag{3.34}$$

where, i is ranges from 0 to n = 1, k represent the force constant used in biasing potential, ξ is collective variable indicating instantaneous position along which the free energy is estimated, and ξ_i is the equilibrium position of the biasing potential. The

weighted histogram analysis method (WHAM) is utilized in finding the probability of biased potential for each window of the molecular system. The free energy profile of hemoglobin protein along a reaction coordinate is obtained from umbrella sampling. Twelve windows are generated each with a separation from the previous window by 1 Å . The initial COM separation was 24.3 Å in between the α and β chain of the dimer. The beta chain is translated away from alpha along the direction of reaction coordinate.



(a)



(b)

Figure 18: (a) Energy profile versus reaction coordinate and (b) probability distribution of biasing potential versus reaction coordinate.

The figure 18 represents the overlap of any two corresponding windows to get the accurate free energy profile. The value of the force constant remains important. Larger and smaller force constant are avoided so as to obtain suitable overlap of the windows. The unbiased free energy profile is given by the function of bias probability distribution

as,

$$G^U(\xi) = f\left(P^b(\xi)\right) \tag{3.35}$$

where, $G^U(\xi)$ is the un-biasing free energy, $P^b(\xi)$ is probability distribution of bias potential energy and ξ is the reaction coordinate along a specific direction of pulling. All the above steps of methodology are implemented to ensure accuracy in the findings, so that they also agree with the available experimental results.

All the above steps of methodology are implemented to ensure accuracy in the findings, so that they also agree with the available experimental results.

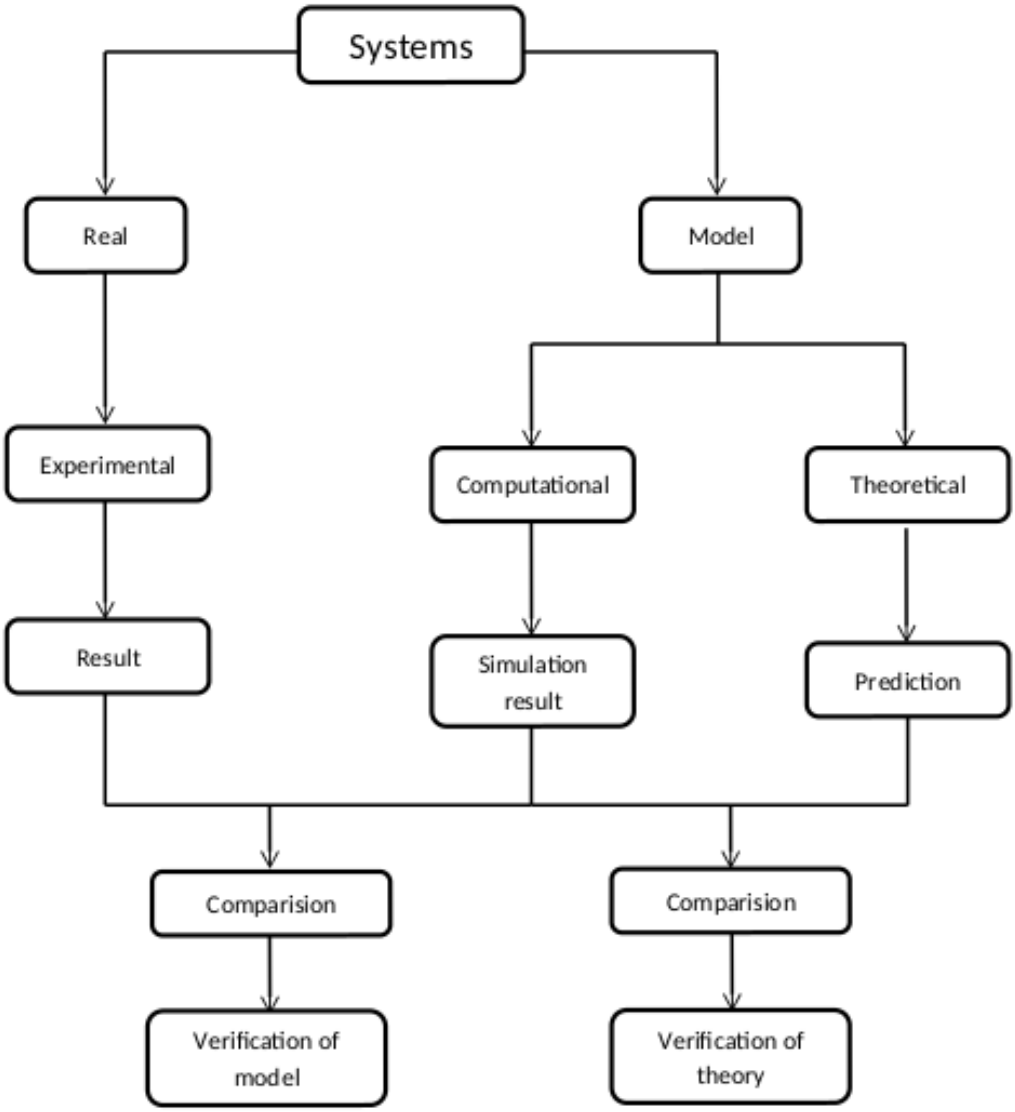


Figure 19: Block diagram showing the relation of computer simulations with experiment and theoretical study.

This Figure 19 is the block diagram which shows the co-relation among the theoretical,

experimental and computational work. Systems are either investigated really through experiment or through theoretically using specific computational model. Both results and findings have certain correlation. Outcomes are used for verifying theory as well as to compare the model with real experiments.

CHAPTER 4

4. RESULTS AND DISCUSSION

4.1 General Consideration

General Consideration In this chapter, we have presented the results and discussion of the molecular dynamics study on sickle and normal hemoglobin protein. In sickle cell hemoglobin, the sixth residue in the beta chain is valine. On the other hand, the sixth residue in the beta chain of normal hemoglobin is glutamic acid. As discussed in subsection 1.1.2, there are four different parts (chains) in a hemoglobin molecule (Figure 3 (a)). The four chains are designated by α_1 , β_1 , α_2 and β_2 chains. $\alpha_1\beta_1$ and $\alpha_2\beta_2$ form dimers of the hemoglobin molecule. These two dimers combine together to form hemoglobin molecule. In this work, we have taken four different systems:

- Normal hemoglobin protein molecule (A3N) as shown in Figure 5 (c).
- Sickle hemoglobin protein molecule (HbS) as shown in Figure 5 (b).
- Dimer of $\alpha_1\beta_1$ ($\alpha_2\beta_2$) of normal hemoglobin as shown in Figure 4 (b).
- Dimer of $\alpha_1\beta_1$ ($\alpha_2\beta_2$) of sickle hemoglobin as shown in Figure 4 (a).

The β chain and α chain of hemoglobin protein carry 146 and 141 active residues respectively on them.

4.2 Results and Discussion

We have studied RMSD of all the systems considered in the present work. In subsection 4.2.1, we have presented the RMSD for the investigation of structural stability. Subsection 4.2.2, 4.2.3, 4.2.4 and 4.2.5 presents the H-bond analysis, hydrophobic interactions, salt bridges, van der Waals (vdW) and electrostatics interactions respectively. Solvent accessible surface area (SASA) analysis is presented in subsection 4.2.6. The results of SMD study for stiffness of hemoglobin protein are presented in subsection 4.2.7. In the subsection 4.2.8, we have presented the analysis of thermal property of hemoglobin dimer whereas, 4.2.9 subsection contains the diffusion study of hemoglobin dimer. Sim-

ilarly, we have presented the binding free energy profile of the dimer structure of sickle and normal hemoglobin protein estimated by umbrella sampling in the section 4.2.10.

4.2.1 Root Mean Square Deviation

We have carried out the systematic study of RMSD versus time for sickled (HbS) and normal hemoglobin (A3N) tetramer to understand their stability. To get more insight, we have separately studied the stability of the alpha chain and beta chain in the tetramer. Further, we have studied the RMSD of their dimer structure. To examine the RMSD, coordinates of the first frame of reference are taken as reference by using the Equation (3.26). Figures 21, 22, 23, 24 and 25 represent the RMSD of both sickle and normal hemoglobin tetramer, alpha chain in the sickle hemoglobin tetramer, alpha chain in the normal tetramer, sickle hemoglobin dimer and normal hemoglobin dimer structures respectively.

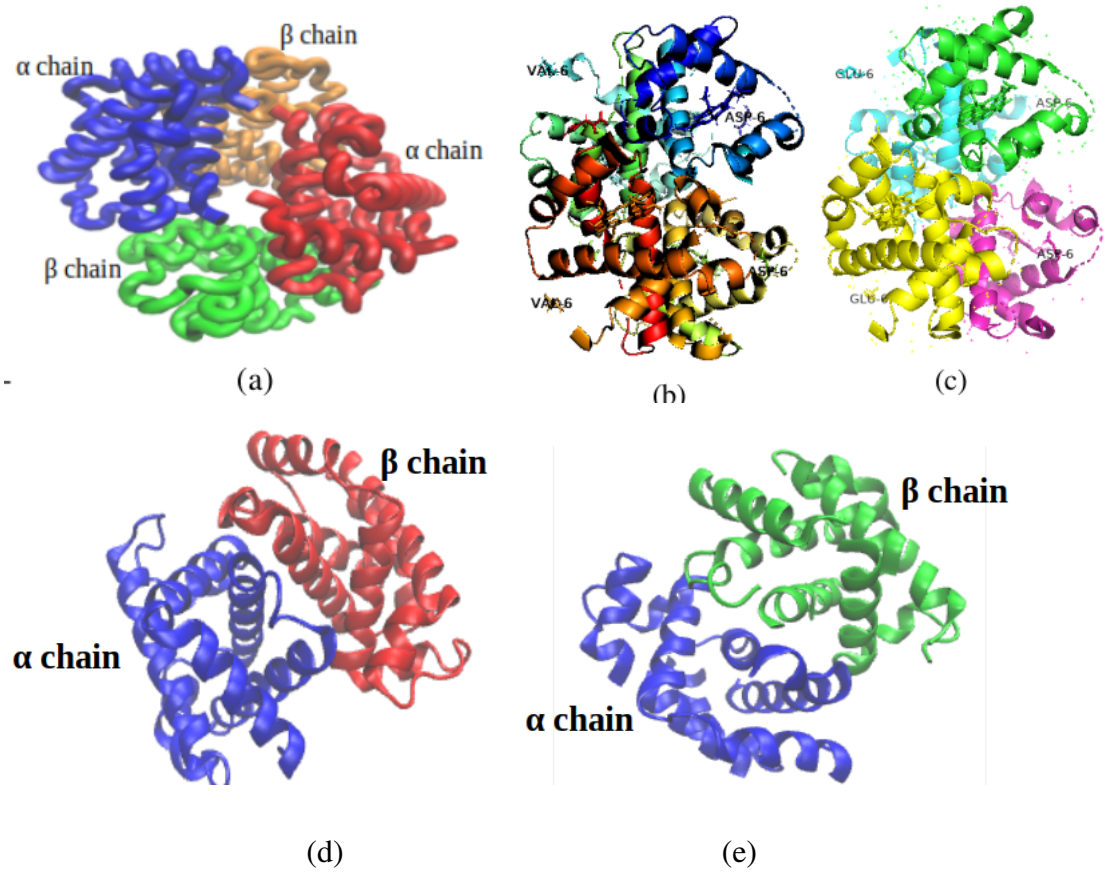


Figure 20: Hemoglobin protein with (a) double - dimer structure, (b) sickle hemoglobin, (c) normal hemoglobin protein, (d) sickle hemoglobin dimer and (e) normal hemoglobin dimer.

The Figure 21(a) is the RMSD graph of both sickle and normal hemoglobin tetramer and Figure 21(b) shows the distribution of RMSD in the tetramer structure of sickle and normal hemoglobin protein.

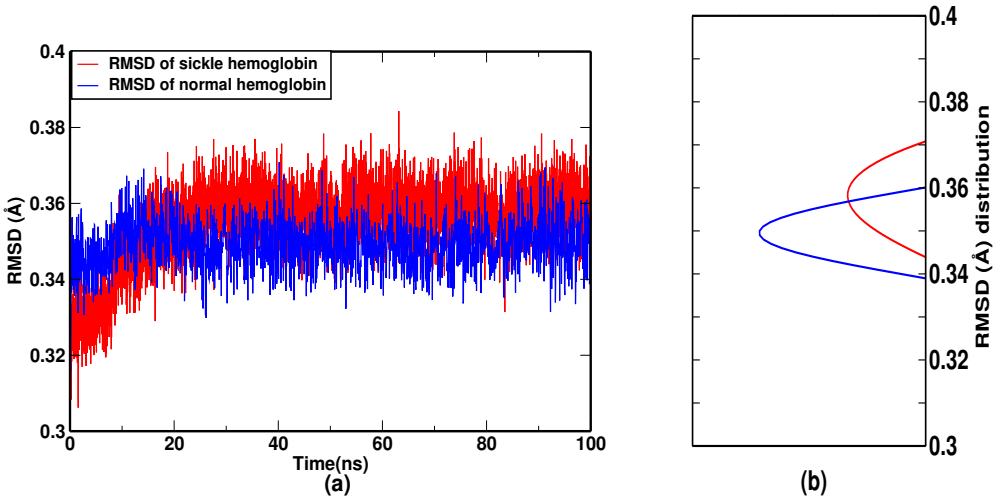


Figure 21: (a) RMSD of sickle and normal hemoglobin (b) distribution of RMSD (Powrel & Adhikari, 2022).

From the Figure 21 (a), it is seen that the RMSD are stable after the simulation time of 20 ns and 25 ns in HbS and 3 respectively. The RMSD graph shows their distribution at Figure 21 (b), which shows the distribution of RMSD of sickle and normal hemoglobin protein. It is seen from Figure 21 (b), that the peak of RMSD distribution lies at 0.36 Å and 0.35 Å respectively for HbS and A3N. This means even though the RMSD fluctuates between 0.32 Å to 0.38 Å in sickle and 0.33 Å to 0.37 Å in normal hemoglobin protein, most of the time they preferred the RMSD values of 0.36 Å and 0.35 Å respectively for HbS and A3N. It shows that the normal hemoglobin molecule is comparatively stable than sickle however, the difference is not significant. This variation may be obtained due to effects of mutation in sickle hemoglobin.

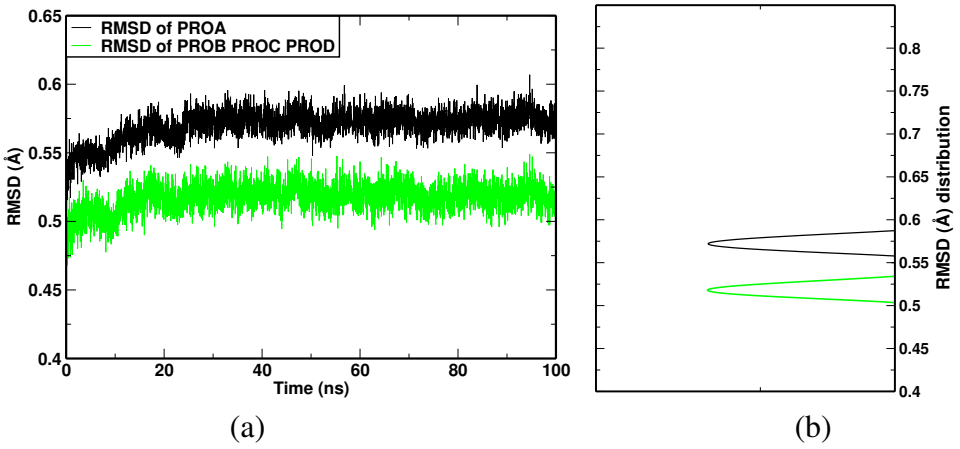
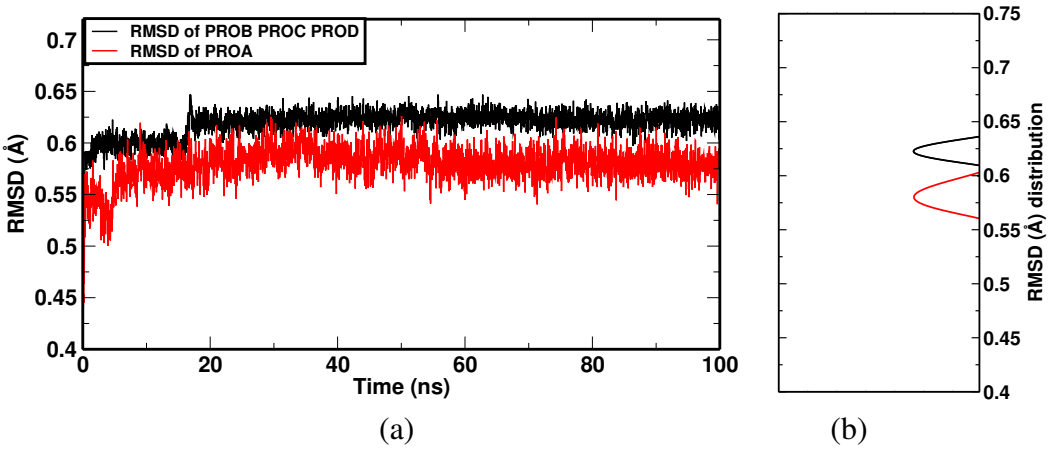


Figure 22: RMSD of (a) sickle hemoglobin and (b) RMSD distribution.

Figure 22 (a) and 22 (b) are the RMSD plots of PROA and PROB PROC PROD of sickle hemoglobin and RMSD distribution respectively. The graph shows that the peak value

of RMSD of chain is at 0.57 Å (Figure 22 (b)). In sickle hemoglobin, most of the time the RMSD of PROA is fluctuates between 0.54 Å and 0.60 Å . Whereas for PROB PROC PROD, it fluctuates within 0.47 Å and 0.57 Å (Figure 22 (a)). The PROA and PROB PROC PROD of sickle hemoglobin protein becomes stable after 20 ns of simulation run.



(a) (b)

Figure 23: (a) RMSD of sickle and normal hemoglobin (b) distribution of RMSD.

Figure 23 (a) and Figure 23 (b) are the RMSD plots of PROA and PROB PROC PROD of normal hemoglobin protein and their RMSD distribution. The RMSD of alpha chain (PROA) in normal is deviated around 0.58 Å and is less than of PROA in sickle.

In normal hemoglobin protein we have seen that (Figure 23), most of the time the RMSD of PROA is fluctuates between 0.50 Å and 0.60 Å . Whereas for PROB PROC PROD, it fluctuates within 0.57 Å and 0.65 Å . PROA of the normal hemoglobin protein is deviated around 0.61 Å and is higher than of sickle protein. RMSD of the PROB PROC PROD, jointly in normal hemoglobin is at 0.63 Å . It shows the higher deviation of chain alpha of sickle than that of chain alpha of normal hemoglobin protein.

We have also estimated the RMSD of both the sickle (Figure 24) and normal hemoglobin dimer (Figure 25) protein of the systems considered for the dimer analysis. In (Figure 24 (a) and (b)) we have shown the RMSD versus time graph of sickle hemoglobin dimer protein and their RMSD distribution. We can see that the structure getting stability after 60 ns of simulation time. However, the chain B is getting stability faster than of chain A as well as sickle dimer protein.

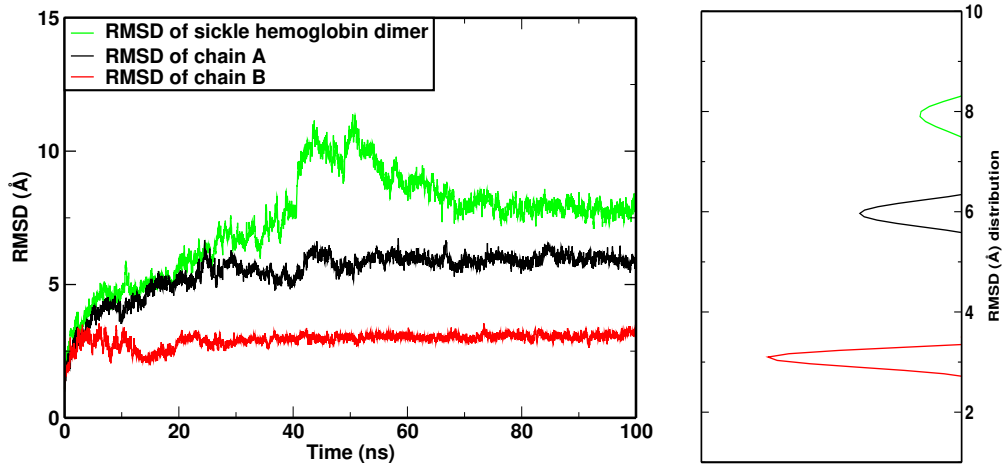


Figure 24: a) RMSD of sickle hemoglobin dimer and b) distribution of RMSD (Powrel et al., 2022)

It is seen that the chain B (β) becomes stable at first and then chain A (α). The whole protein takes comparatively longer time to become stable then of α and β chain.

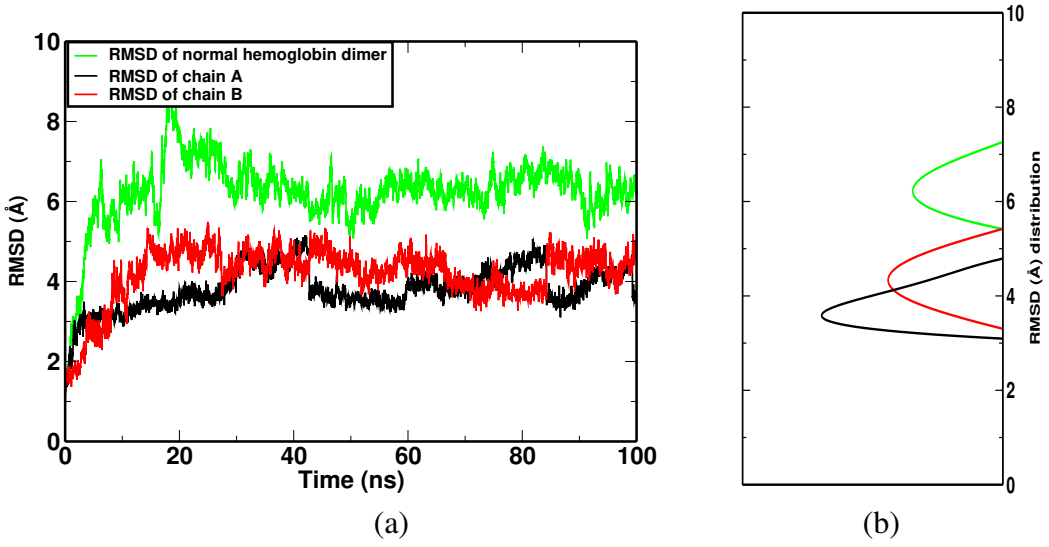


Figure 25: a) RMSD of normal hemoglobin dimer and b) distribution of RMSD (Powrel et al., 2022)

Figure 25 (a) is the RMSD versus time graph of normal hemoglobin protein dimer. For normal hemoglobin, we can see that the PROA, PROB are fluctuating more than as they appears in sickle tetramer (Figure 24 (a)). The whole dimer protein of normal hemoglobin becomes stable after 40 ns. It is as faster as that of sickle protein dimer. Figure 24 (a) and Figure 25 (a) shows that the root mean square deviation (RMSD) of PROA is higher than of PROB and whole protein in both the dimer, i.e., PROA is more flexible and PROB is less flexible or more consistent. Figure 24 (b) and 25 (b) are giving the distribution of RMSD of dimer HbS and A3N respectively. The Figure 24 (a) shows that the RMSD of sickle protein is fluctuating between 5 Å and 9 Å . However, in normal

hemoglobin (Figure 24 (a)) the RMSD is fluctuating between 4.5 Å and 8.5 Å . Most of the time the peak value of RMSD of sickle lies at 8 Å and that of the normal dimer is at 6 Å , i.e., the deviation of RMSD in sickle is slightly higher than that of normal hemoglobin. RMSD analysis of sickle and normal dimer protein strictly saying that the sickle is comparatively unstable. Even though, the two complexes with their entire molecules under dissolve condition in water are stable in structure. Thus the structures are further implemented in finding their free energy to evaluate the binding affinity and structural variation during the molecular dynamics simulations.

4.2.2 Hydrogen Bond Interactions

Hydrogen bonding plays crucial role to form the secondary and higher order structures of protein and nucleic acid. This bonding mechanism not only contributes in intra molecular formation but also the formation of the molecular complex of two or more molecules (Eaton, 2002) (Eaton et al., 2007). The folding and unfolding of protein structure are the results of hydrogen bonds between the backbone-backbone and side chain-backbone of the molecules. Biomedical activity and bio-mechanisms are influenced by hydrogen bonding of the atoms in the molecule. Hydrogen bonds are also responsible for the manifestation of the three-dimensional structure of folded proteins and takes an important role in their biochemical changes. In this study, we have estimated the number of hydrogen bonds and the hydrogen bond occupancy between different chains of hemoglobin molecules. To study the hydrogen bond interaction, the residues involved in the hydrogen bonding as well as the number of hydrogen bonds in the α chain of HbS and A3N, we used the cut-off distance parameters of 0.35 nm and 30^0 for length and angle respectively.

Hydrogen Bonds in Sickle and Normal Hemoglobin

Figure 26 (a) presents the variation in the occurrence of number of hydrogen bonds in the residues of sickle and normal hemoglobin. From the Figure 26 (a), we can see that the highest number of hydrogen bonds in sickle hemoglobin is found to be 161 and the lowest number of hydrogen bonds is 132. Similarly, in the case of normal hemoglobin, the highest number of hydrogen bonds is 154 and lowest number of hydrogen bonds is 129. The average number of hydrogen bonds in sickle hemoglobin is 147 whereas that in normal is 142.

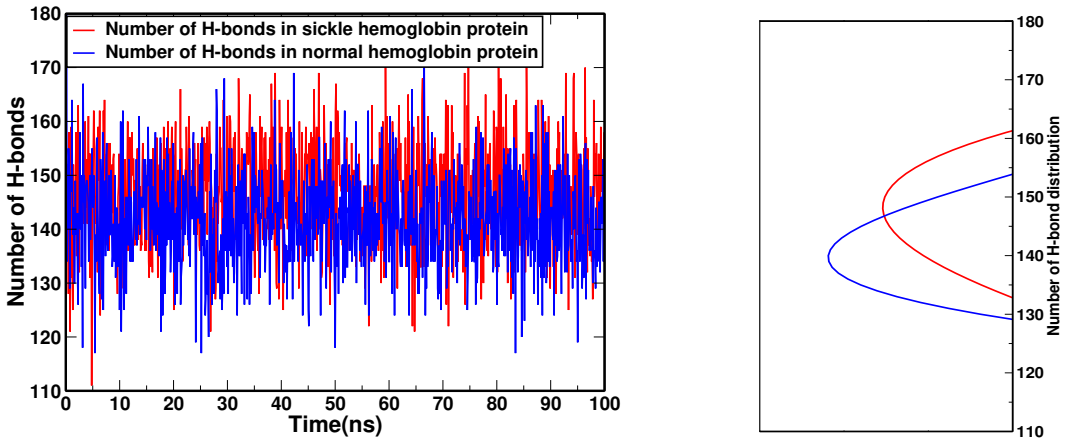


Figure 26: (a) Time evolution of hydrogen bonds in sickle and normal hemoglobin protein and (b) distribution of the number of hydrogen bonds in sickle and normal hemoglobin protein tetramer (Powrel & Adhikari, 2022)

Figure 26 (b) shows the distribution of number of hydrogen bonds in sickle and normal hemoglobin protein. The peak values in the distribution curves are 147 and 142 in HbS and A3N also obtained in Figure 26 (a). The peak value of the distribution of the number of hydrogen bond interactions is greater in sickle than that of normal hemoglobin as in Figure 26 (b).

We have also studied the occupancy percentage of the hydrogen bonds present in the system. Figure fig:hbondoccbothbeta presents the occupancy of the residue pairs having hydrogen bonds in sickle (red color) and normal hemoglobin (black color) protein. There were 428 pairs of residues involved in formation of hydrogen bonds in the sickle protein. Out of them 29 pairs have shown occupancy above 100 %. The most stable H-bond was of SER49-ASP47 with an occupancy of 310 %.

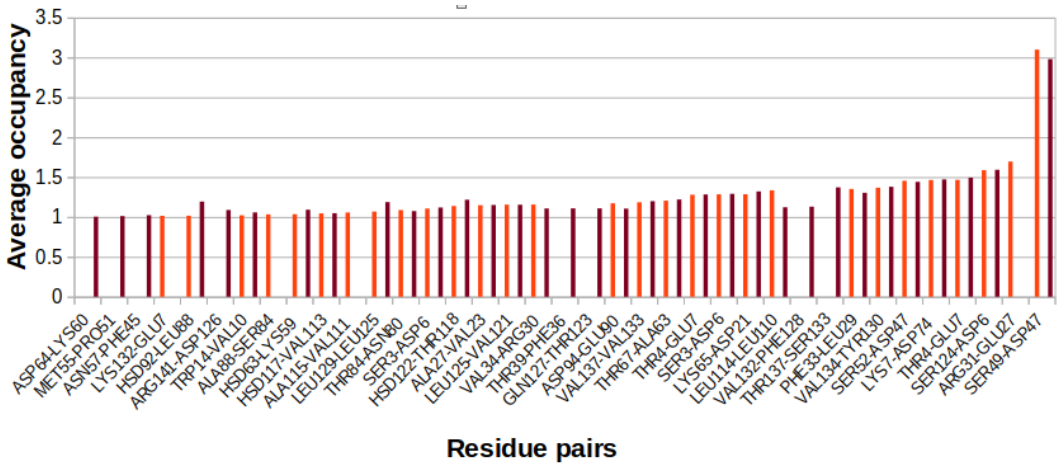


Figure 27: Average occupancy of H-bond versus pairs of residues involved in sickle and normal hemoglobin tetramer (color online) (Powrel & Adhikari, 2022).

However in normal hemoglobin, there were only 378 pairs of residues which can form hydrogen bonds. Among them, 33 pairs have shown the occupancy above 100 %. The SER49-ASP47 residue have shown the maximum occupancy of 298 %. Potential nucleophile residues of lysine (LYS) are most abundantly found in hydrogen bonding. The SER49-ASP47 residue pair is found most consistent throughout the simulation (Figure 27) in both the protein.

Hydrogen Bonds in α Chains of Sickle and Normal Hemoglobin Protein

As mentioned in the introduction section, a hemoglobin protein contains four different peptide chains, namely α_1 (PROA), β_1 (PROB), α_2 (PROC) and β_2 (PROD). To analyze hydrogen bonds in such proteins, we have studied number of hydrogen bonds in α_1 (PROA) chain and other three remaining chains (PROB PROC PROD) separately. This gives the idea about how an individual chain, out of four chains of the protein, contribute to the hydrogen bonding in hemoglobin protein.

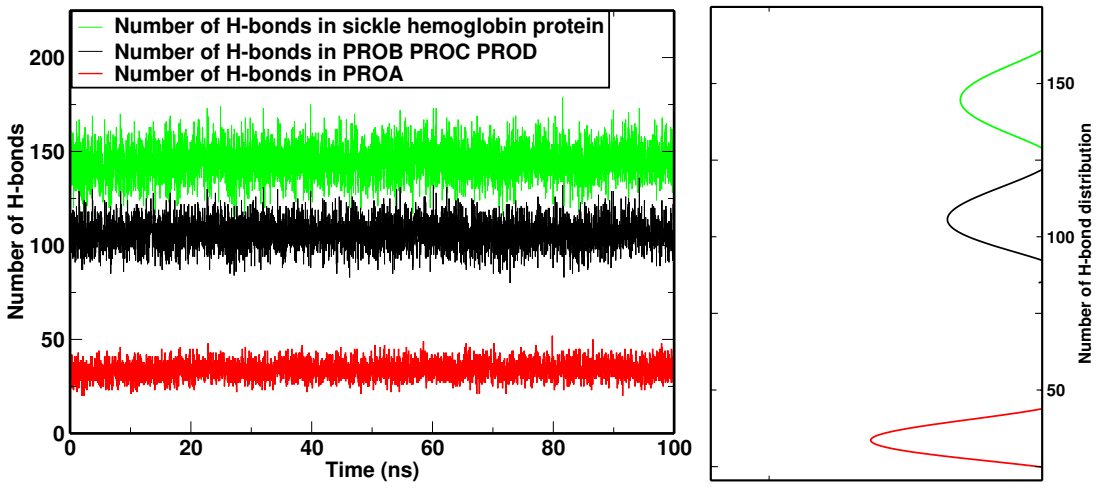


Figure 28: (a) Number of hydrogen bonds in the sub-units and (b) distribution of the hydrogen bonds in sickle Hb tetramer.

Figure 28 (a) depicts the number of hydrogen bonds in α_1 chain and that in other three chains (PROB PROC PROD) in sickle hemoglobin. As usual the number of hydrogen bonds fluctuates between 25 and 44 in PROA and 92 and 122 in PROB PROC PROD of sickle hemoglobin protein as presented in Figure 28 (b). Figure 28 (a) and (b) are the number of the formation of H-bonds and the distribution of H-bonds in sickle hemoglobin protein respectively. The graph shows that the PROA (α chain), PROB PROC PROD and sickle hemoglobin protein form 35, 112 and 147 H-bonds (Figure 28 (b)) respectively.

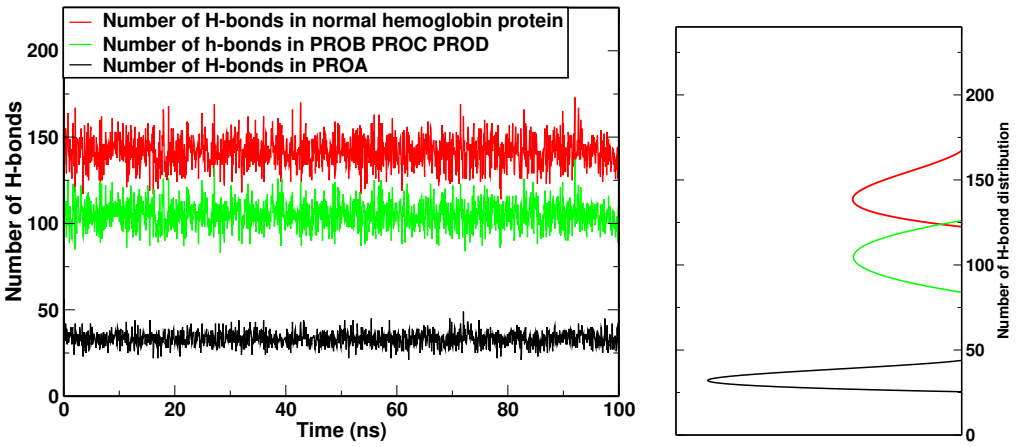
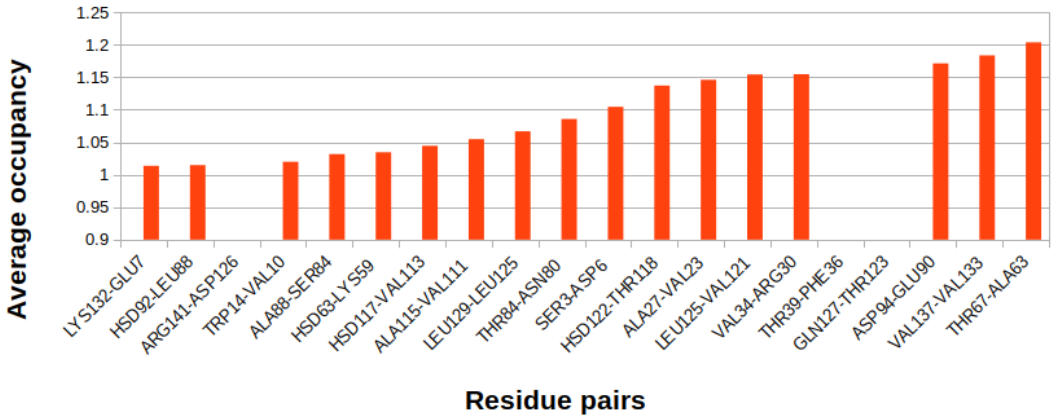


Figure 29: (a) Number of hydrogen bonds in the sub-units and (b) distribution of the number of hydrogen bonds in normal Hb.

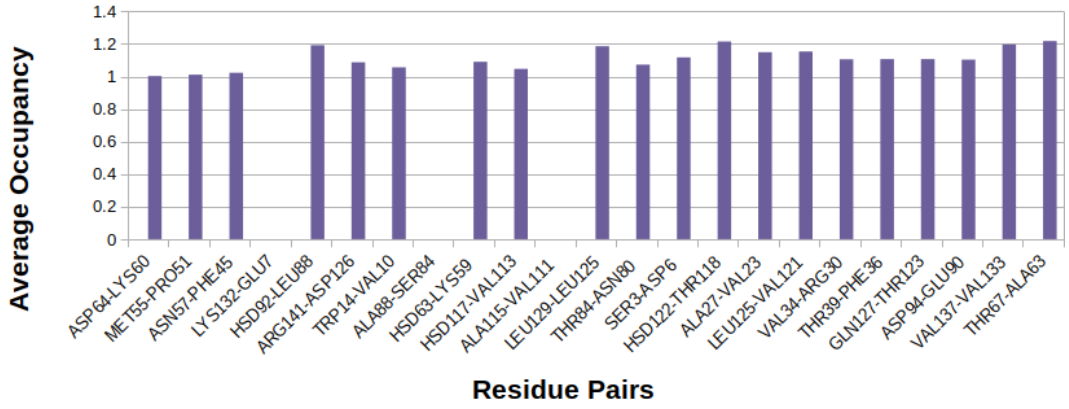
Figure 29 (a) depicts the number of hydrogen bonds in α_1 chain and that in other three chains (PROB PROC PROD) in normal hemoglobin. As usual the number of hydrogen bonds fluctuates between 26 and 42 in PROA and 100 and 126 in PRO BCD of normal hemoglobin protein as presented in Figure 29 (b).

We have also estimated the H-bonds in normal hemoglobin, PROA and PROB PROC PROD of normal hemoglobin protein (Figure 29 (a)). It shows that the PROA, PROB PROC PROD and normal hemoglobin protein form 35, 107 and 142 H-bonds respectively. From Figures 28 and 29, it is seen that the normal hemoglobin protein as well as PROA of normal hemoglobin, comparatively form a lower number of H-bonds than in sickle hemoglobin protein.

The active residue pairs involved and their average occupancy percentage in the hydrogen bond formation are also studied for sickle (Figure 30 (a)) and normal hemoglobin (Figure 30 (b)). Some of the residues involved in hydrogen bond formation in sickle were disappearing in normal hemoglobin and a few of new residues were involved in normal Hb than that in sickle hemoglobin as in Figure 30.



(a)



(b)

Figure 30: Average occupancy of hydrogen bond versus residue pairs involved in (a) sickle and (b) normal hemoglobin tetramer during alpha SMD.

Figure 30 (a) shows that the residue pairs ARG141-ASP126, THR39-PHE36 and GLN127-THR123 from sickle are disappearing which were found active to form hydrogen bonds in normal hemoglobin (Figure 30 (b)). Similarly, some other residues like LYS132-GLU7, ALA88-SER64 and ALA115-VAL111 are not found in normal which have been actively participating in sickle for hydrogen bond formation (Figure 30 (b)).

Hydrogen Bonding in Dimer Hemoglobin

A hemoglobin dimer is the active molecular complex of α and the β peptide chains (Figure 4). The hydrogen bonds of dimer have certain important role in providing the elasticity to the whole structure of proteins. It also helps in the formation of single helical structure as well as in the interaction among the different chains. They also contributes to the intra and inter-molecular complex formation. Each of such protein chain folds by means of intra protein hydrogen bonds in them.

This study estimates the instantaneous number of intra-chain and inter-chain hydrogen bonding of the sickle and normal hemoglobin dimers. The residues involved in the

hydrogen bonding of α chain and β chain in both complexes are identified. For this, we used the cut-off distances of 3.2 Å and cut-off angle of 30^0. Figure 31 and 32 shows the number of hydrogen bonds in the residues of the dimer proteins. We have also focused on determining the occupancy percentage of hydrogen bonding in the residue pairs.

Initially, there were more number of hydrogen bonds in sickle dimer protein of around 40, and then continuously reduced to a constant value as shown in Figure 31. Twenty four of the H-bonds are constantly present in the sickle dimer giving elastic properties to the protein on average (Figure 32). Figure 31 (a) also depicts the number of hydrogen bonds in α_1 (PROA) chain and β_1 chain (PROB). The number of hydrogen bonds mostly fluctuates between 15 and 50 in sickle protein, 3 and 17 in PROA and 4 and 16 in PROB of sickle hemoglobin dimer protein as presented in Figure 31 (b). The average number of hydrogen bonds in sickle dimer protein, PROA and PROB are found to be 24, 10 and 10 respectively.

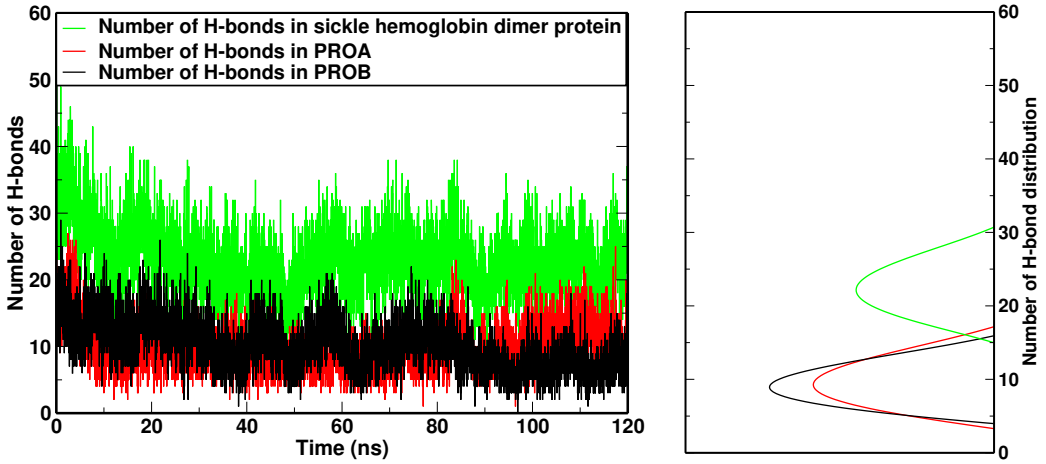


Figure 31: (a) H-bond in sickle hemoglobin dimer and (b) average number of the H-bonds.

During the simulation of the dimer of sickle hemoglobin, it was seen that the number of H-bond formation fluctuates throughout the simulation. Among the two protein chains in the dimer system, beta chain and alpha chain have shown the same number of H-bond contribution in the stability of the sickle dimer structure (Figure 31 (b)).

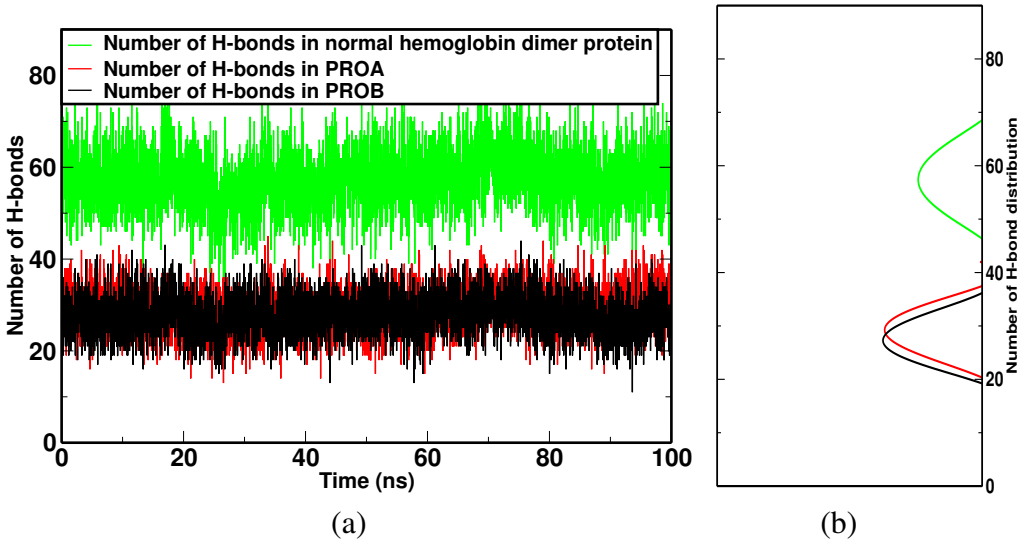


Figure 32: (a) Number of H-bond in normal hemoglobin dimer and (b) distribution of number of the H-bonds (Powrel et al., 2022)(color online).

The normal hemoglobin dimer, however, shows 60 hydrogen bonds in most of the time to the dimer protein (Figure 32 (a)). Figure 32 (a) also depicts the number of hydrogen bonds in α_1 (PROA) chain and β_1 chain (PROB). As usual the number of hydrogen bonds mostly fluctuates between 38 and 78 in normal dimer, 20 and 37 in PROA and 19 and 36 in PROB of normal hemoglobin dimer protein as presented in Figure 32 (b). The average number of hydrogen bonds in normal dimer protein, PROA and PROB are found to be 60, 30 and 29 respectively. It is also seen from the H-bond estimation that the normal hemoglobin dimer is more elastic than the sickle dimer.

The number of inter chain H-bond between PROA and PROB of the sickle dimer is observed as 0.47. Whereas, the number of H-bond between PROA and PROB of normal hemoglobin is found to be 0.70. i.e., a weak presence of hydrogen bond was seen between chain α and β in the both dimer. Even the presence of H-bond between α and β in normal is higher than of sickle dimer.

Figure 33 and Figure 34 are the averages of the hydrogen bond occupancy of residue pairs in sickle and normal hemoglobin dimer respectively.

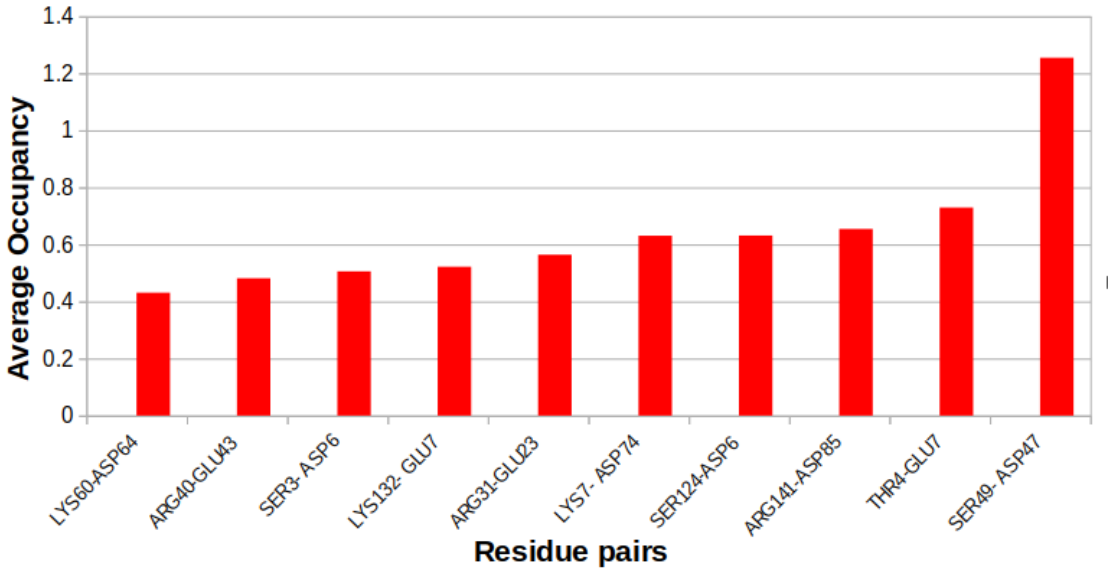


Figure 33: Average occupancy of hydrogen bond versus residue pairs involved in (only above 40 %) sickle hemoglobin dimer (Powrel et al., 2022).

Figure 33 represents the occupancy of hydrogen bond above 40 % in the sickle dimer protein. A maximum occupancy of 1.26 % have been estimated for the residue pair of SER49-ASP47 in sickle dimer. Figure 34 represents the occupancy of hydrogen bond above 40 % in the normal hemoglobin dimer protein. It shows that the SER49-ASP47 have maximum occupancy of 1.38 % throughout the simulation.

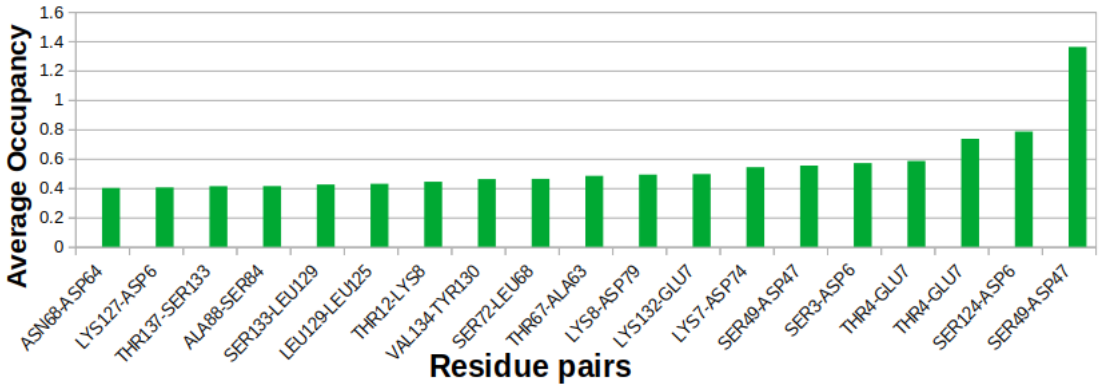


Figure 34: Average occupancy of hydrogen bond versus residue pairs involved in (only above 40 %) normal hemoglobin dimer (Powrel et al., 2022).

The H-bond occupancy in normal hemoglobin dimer (Figure 34) shows 19 residue pairs having occupancy above 40 %. More residue pairs are involved in the H-bonds of normal hemoglobin protein dimer to provide the strength of elasticity in the protein. It is seen that the SER124-ASP6 residue which has active participation in hydrogen bonding of normal was not seen in sickle hemoglobin.

This study of hydrogen bond interactions in hemoglobin proteins as well as in dimer structure of hemoglobin protein shows that the dimers of normal hemoglobin protein

are much stronger than that of sickle hemoglobin dimer as it was bounded with more hydrogen bonds. However, in tetramer structure the sickle hemoglobin is strongly bonded by hydrogen bonds than that in normal hemoglobin.

The study further shows that the number of hydrogen bonds in sickle hemoglobin tetramer structure increases to 147 which were only 24 in sickle dimer. Similarly, there were 60 hydrogen bonds in the normal hemoglobin dimer which were found 142 in normal hemoglobin tetramer. In both cases the number of hydrogen bonds found more than the double of the number of hydrogen bonds in their dimer structure. The results also depicts the stronger occupancy of H-bond in sickle hemoglobin than that of normal hemoglobin protein.

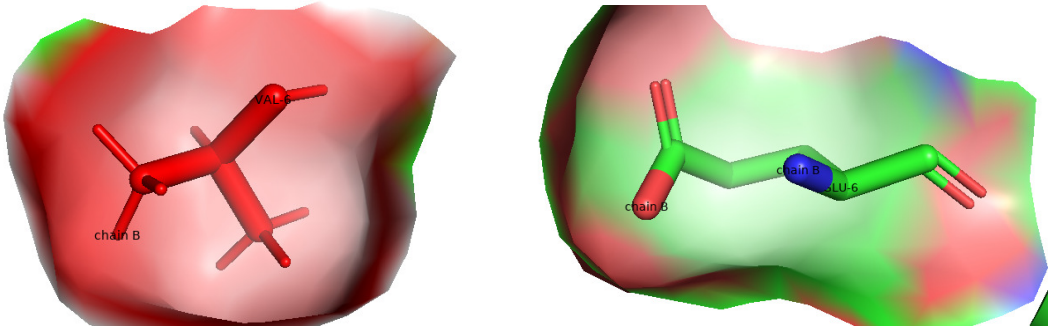
4.2.3 Hydrophobic Interactions

The hydrophobic potential is another important interaction associated with the build-up of protein structure (Tina et al., 2007). The strength of the hydrophobic interaction of the molecular surface comes from the number of hydrophobic residues involved in the protein structure using PIC.mbu software (Tina et al., 2007). It brings changes in the overall shape of the molecule so that the inter- residue interaction enhances the structure of the molecules to form. The hydrophobic surfaces try to reduce the overall surface size of molecules. The inter-residue interaction and folding of protein occurs in the system of molecules due to the hydrophobic surface interaction of the residues at the surface.

We have estimated the number of hydrophobic interactions and residues involved in the sickle and normal hemoglobin, alpha chain in the sickle and normal hemoglobin and in dimer structure. We have also studied the hydrophobic interactions of different frames (Tina et al., 2007) (Tina et al., 2007).

Hydrophobic Interactions in the Sickle and Normal Hemoglobin Protein

We have studied the protein-protein hydrophobic interactions in the sickle and normal hemoglobin. Also, in the case of the dimer, we have estimated the number of hydrophobic residues involved in the complexes. Figure 35 (a) and Figure 35 (b) are the hydrophobic and hydrophilic surfaces of 6th position of valine and glutamic acid buildup within the beta chain of sickle and normal hemoglobin protein. We have seen 491 and 303 residue pairs involved in the hydrophobic interactions of the sickle and normal hemoglobin respectively. We have also examined the inter-chain hydrophobic interactions inside them.



(a) (b)

Figure 35: a) Hydrophobic nature of 6^{th} valine in in β chain of sickle and b) hydrophilic nature of 6^{th} glutamic residues in β chain of normal hemoglobin protein.

There were 32 sets of hydrophobic residues actively participating in between protein-protein chains (Powrel & Adhikari, 2022) of sickle hemoglobin, out of which 16 residues involved only from the beta chain of sickle hemoglobin. The remaining 16 pairs of residues were seen within the other three chains (i.e., of two alpha and one beta chains). All other 459 pairs of residues pairs were seen as intra-protein hydrophobic interactions of sickle hemoglobin structure. Leucine (LEU) residue was found mostly (173 sets) in contributing to the hydrophobic interaction in HbS. Out of 303 residue pairs of normal hemoglobin (Powrel & Adhikari, 2022), 271 pairs are intra-protein hydrophobic interactions. Among them eight of the residues, LEU, VAL, PHE, ALA, TRP, PRO, TYR and MET form 89, 55, 40, 35, 16, 15, 11 and 10 sets of hydrophobic interactions respectively. The remaining 32 sets of residue pairs are still carrying the protein-protein hydrophobic interaction in typical hemoglobin. The chain β mostly interacted with hydrophobic potential in both proteins.

Hydrophobic Interaction in Alpha Chain of Sickle and Normal Hemoglobin

There were 105 intra-protein hydrophobic interactions actively participating in residue pairs of alpha chain (PROA) in sickle hemoglobin. However, alpha chain of normal hemoglobin have shown only 95 active hydrophobic interactions among the residue pairs in them.

Hydrophobic Interaction in Dimer

The hydrophobic interaction provides the inter-residue activity in the dimer protein in water (Tina et al., 2007). The systems of sickle and normal hemoglobin dimer in five different time frames are picked up as the frame 500, 1500, 2500, 3500 and 4500. From these outlines, we have found the number of residue pairs including in intra-protein hydrophobic interaction of sickle and normal hemoglobin protein dimers (Table 1).

Table 1: The number of hydrophobic sets of sickle and normal hemoglobin dimer protein at five different consecutive time frames (Powrel et al., 2022).

For	Number of residue pairs involved					Average number
Frame	500	1500	2500	3500	4500	-
Sickle Hb dimer	233	235	232	224	227	230
Normal Hb dimer	224	207	207	216	216	214

Hydrophobic interactions inside the sickle and normal hemoglobin protein dimer are studied separately. During simulation more hydrophobic interactions were found in sickle dimer than of ordinary hemoglobin dimer (Table 1). From the study of hydrophobic interaction in sickle and normal hemoglobin, we have found more contribution of hydrophobic interaction in sickle hemoglobin, alpha chain of sickle and dimer of sickle hemoglobin than that in corresponding normal hemoglobin. i.e., the higher contribution of hydrophobic interaction were seen in sickle hemoglobin as well as in their dimer than that of corresponding normal hemoglobin structures.

It is also seen that, during tetramer formation of sickle hemoglobin some of the new residue pairs gains hydrophobic interaction which were not seen earlier in their dimer structure. However, in normal hemoglobin some of the residue looses their hydrophobic interactions which were present in their dimer structure.

4.2.4 Salt Bridge Interactions

Salt bridge is the interaction of amino acids with oppositely charged ions which prevents the diffusion of solution (Marcus & Hefter, 2006). They are also important to form the stable structure of the proteins. We have studied the detail of the salt bridge interactions among the residue pairs in the protein chains of both sickle and normal hemoglobin proteins. We have also compared the salt bridges of hemoglobin proteins with that found in the dimer hemoglobin structure. We have also estimated the occupancy of residue pairs involved in the formation of salt bridges and their corresponding occupancy percentage in both sickle and normal hemoglobin, alpha chain and sickle and normal hemoglobin dimer respectively in Figures 36, 37 (a), 37 (b), 38 and 39.

Salt Bridges in Sickle and Normal Hemoglobin

Figure 36 represents the average of salt bridge occupancy in sickle and normal hemoglobin protein. We have analyzed the salt bridges in the different chains in the proteins. The result indicated that the residue pairs of ASP94-LYS95, ASP94-LYS95 and GLU43-ARG92 have the occupancy percentage above 5 in the HbS (Figure 36 (red colored bar graph) protein. There was GLU43-ARG92 residue pair having the highest contribution of salt bridge interaction to the HbS protein with occupancy of 5.80.

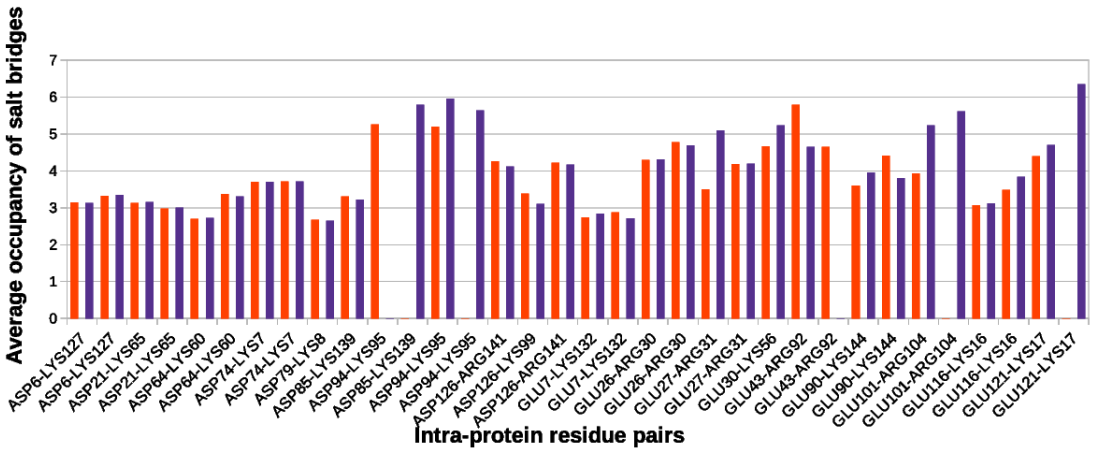
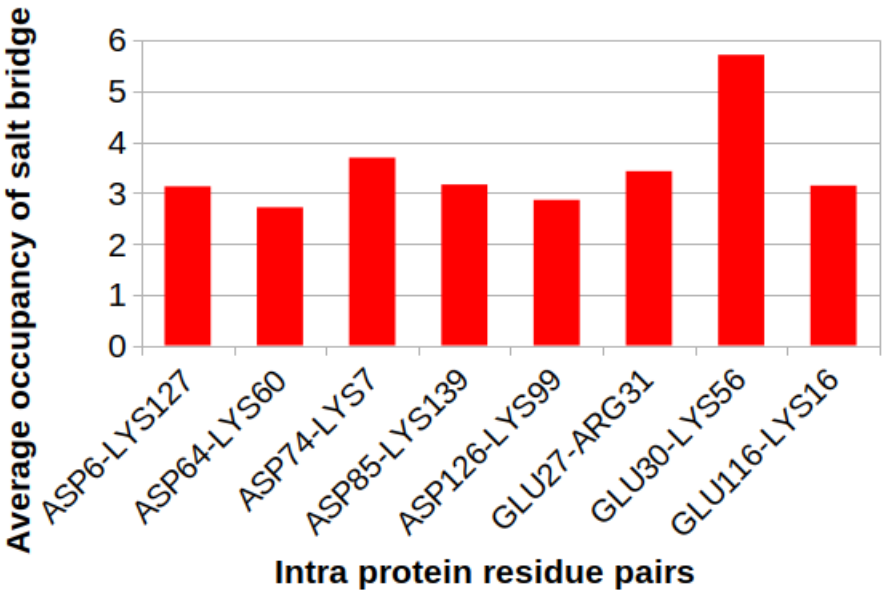


Figure 36: Graph of the average salt bridge occupancy of different residue pairs in the sickle and normal hemoglobin protein (color online).

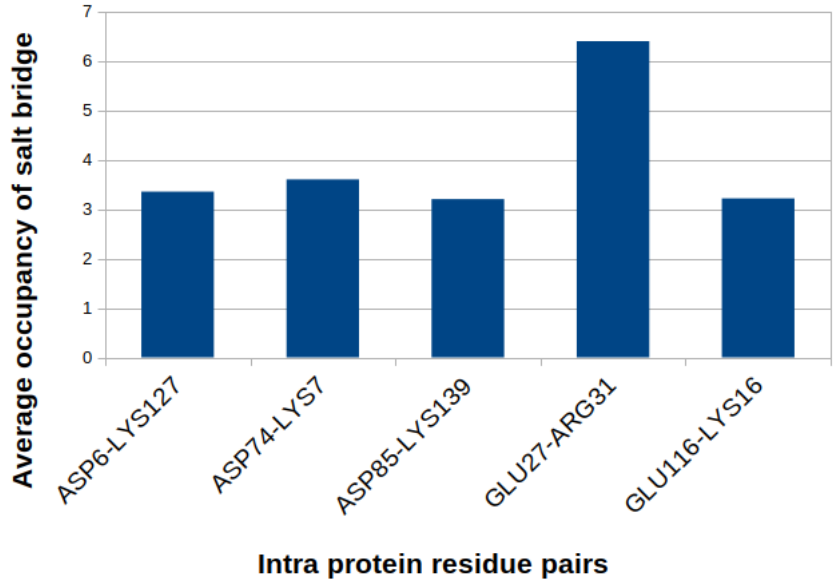
There were 30 salt bridges effectively appearing in sickle hemoglobin protein. However, for the normal hemoglobin, we have found 32 sets of residue pairs with dynamic salt bridges (Figure 36 (blue colored bar chart)). Among them, 8 sets of residue pairs GLU121-LYS17, ASP94-LYS95, ASP85-LYS139, ASP94-LYS95, GLU101-ARG104, GLU101-ARG104, GLU30-LYS56 and ASP85-LYS139 have occupancy above 5. GLU121-LYS17 residue pairs contribute highest salt bridge occupancy in normal hemoglobin protein (A3N) with the average occupancy of 6.30. In all chains, ASP and GLU residues were found to be most abundant in salt bridge formation.

Salt Bridges in Alpha Chain of Hemoglobin Protein

We have also estimated the residue pairs involved in salt bridge formation and their occupancy in the chain α (PROA) of sickle and normal hemoglobin protein structures.



(a)



(b)

Figure 37: Salt bridges in PROA of (a) sickle and (b) normal hemoglobin tetramer.

Figure 37 (a) and 37 (b) have presented the average numbers of salt bridges and the residue pairs involved in the chain of sickle and normal hemoglobin protein respectively. Eight (8) pairs of residues were found in the salt bridge interaction of alpha chain of sickle hemoglobin. However, the alpha chain of the normal hemoglobin structure have shown the involvement of only 5 residue pairs in their salt bridge interaction.

More residue pairs were involved in the salt bridges of the alpha chain of sickle than that of the alpha chain of normal hemoglobin protein. This study have also shown that the GLU30-LYS56 and GLU27-ARG31 residue pairs possesses maximum occupancy in

sickle and normal hemoglobin respectively.

Salt Bridges in Hemoglobin Dimer

We have also studied the detail of the salt bridges of sickle and normal hemoglobin dimer. Figure 38 and Figure 39 represent the average of salt bridges occupancy in sickle and normal hemoglobin dimer respectively.

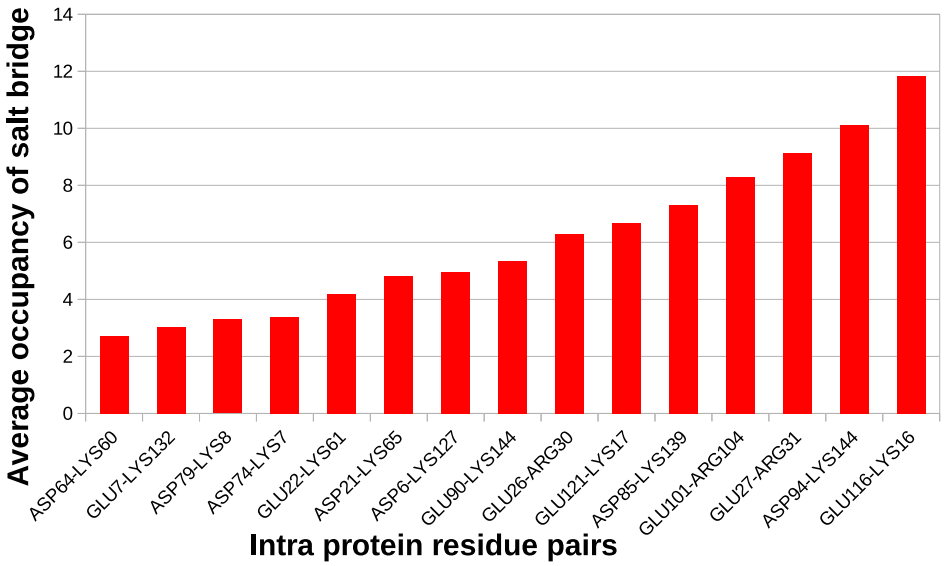


Figure 38: Salt bridge occupancy in sickle hemoglobin dimer (Powrel et al., 2022).

In sickle hemoglobin dimer, there were 15 pairs of active intra protein salt bridges, among them 6 pairs in chain A and 9 pairs in chain B (Figure 38). One salt bridge (GLU101-LYS99) was found in between the chains α and β in the sickle dimer. The ASP94-LYS144 and GLU116-LYS16 of sickle hemoglobin dimer have shown the occupancy above 10 % (Figure 38) in sickle hemoglobin dimer

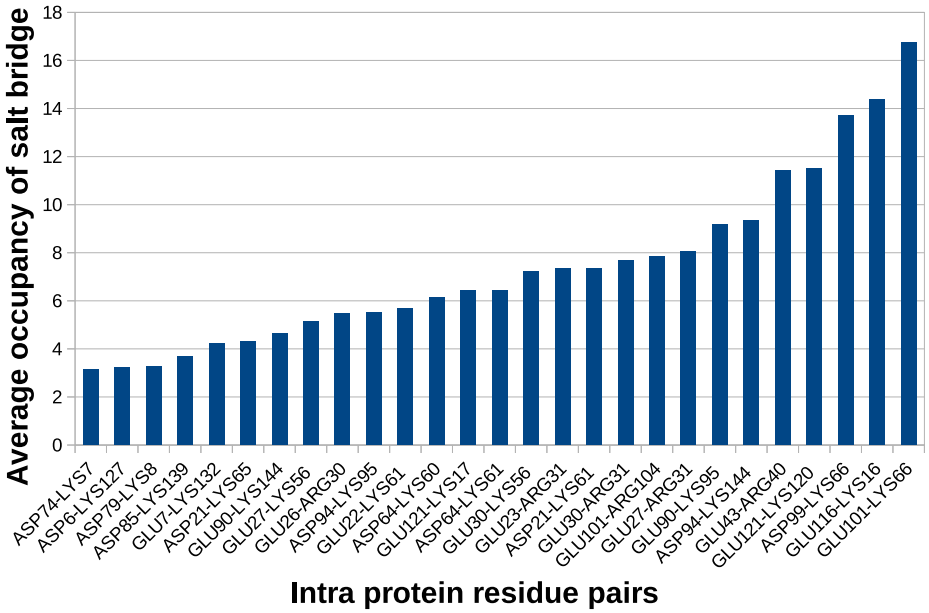


Figure 39: Salt bridge occupancy in normal hemoglobin dimer (Powrel et al., 2022).

However, in normal hemoglobin protein dimer, there were 27 pairs of intra protein salt bridges. Among them, 11 salt bridges were found in alpha chain, 16 salt bridges in beta chain of normal dimer. Additionally, 7 inter chain salt bridges were also found between alpha and beta chain of normal hemoglobin dimer. ASP-LYS, GLU-LYS, GLU-ARG and ASP-ARG pairs were abundantly found in salt bridges of normal dimer (Figure Figure 39). The GLU43-ARG40, GLU121-LYS120 , ASP99-LYS66, GLU116-LYS16 and GLU101-LYS66 of normal hemoglobin have shown the occupancy above 10 % (Figure 39) in normal hemoglobin dimer.

This salt bridge study of hemoglobin structure and dimer structure of sickle and normal hemoglobin shows that GLU, ASP and LYS are mainly contributing in salt bridge formation of hemoglobin protein. Also, more residue pairs are involved in salt bridge interaction of the normal hemoglobin as well as dimer. However, the chain alpha of normal hemoglobin contributes less to the salt bridge interaction than that by the alpha chain of the sickle hemoglobin. This result favors the stronger binding in normal hemoglobin.

4.2.5 vdW and Electrostatic Interactions

Electrostatic and van der Waals interaction appears due to the ions present in the molecules. However, in molecules, the atoms share their valence electrons and the electron density is shifted due to different electro-negativity of the atoms. These different charge densities can be mapped into partial point charges which are represented by the Coloumb charge. They have an important role in the intra as well as in inter-molecular

bindings of proteins. Some amino acids have positive or negative and non-zero charges. In addition to these, atoms in the molecules contain partial charges that are not zero. Such a molecule frequently takes the form of dipole moments. Electrostatic interactions between these partial or net charges in the molecules are important for binding the chains in protein molecules as well as in the formation of magnetic moments (González-Ruiz & Gohlke, 2006). Estimation of such non-bonded energy could provide the information on the dynamics of protein at about the possible distance at which the folding and unfolding of protein occurs.

We study the vdW and electrostatic interactions in the hemoglobin and their dimer structures. These interaction energies for sickle hemoglobin, normal hemoglobin and dimers of sickle and normal are presented in the Tables 2, 3, 4, 5, 6 and 7.

vdW and Electrostatic Interaction of Sickle, Normal Hemoglobin and Alpha Chain

We have also studied the vdW and electrostatic interactions present in the sickle and normal hemoglobin protein and alpha chain. The interaction energies are tabulated for sickle, normal hemoglobin and alpha chain (Table 2 and 3).

Table 2: The average vdW and electrostatic interaction potential energy in sickle hemoglobin protein and its alpha chain (Powrel et al., 2022).

For	vdW (kcal/mol)	Electrostatic (kcal/mol)	Sum (kcal/mol)
Equilibration HbS	-1663.02	-7023.26	-8686.28
Equilibration α chain	-1141.78	-3913.30	-5055.08

Table 3: The average vdW and electrostatic interaction potential energy in normal hemoglobin protein and its alpha chain (Powrel et al., 2022).

For	vdW (kcal/mol)	Electrostatic (kcal/mol)	Sum (kcal/mol)
Equilibration A3N	-1705.00	-6505.72	-8210.72
Equilibration α chain	-54.91	-1280.81	-1335.72

Table 2 and Table 3 represent the averages of the vdW and electrostatic interaction potential energies in sickle and normal hemoglobin proteins. Comparing the data from Table 2 and Table 3 obvious that the sum of the vdW and electrostatic interaction energy of sickle hemoglobin and its alpha chain is higher than that of corresponding normal hemoglobin. It was also seen that the electrostatic potential energy in both HbS and A3N were much higher than vdW potential energy.

vdW and Electrostatic Interaction of Sickle, Normal hemoglobin and Beta Chain

In this study, we have estimated the inter-molecular electrostatics and vdW energy in sickle and normal hemoglobin proteins. We have also presented the quantitative

estimation of the electrostatics and vdW energy of the beta chain.

Table 4: The average vdW and electrostatic interaction potential energy in sickle hemoglobin protein and its beta chain (Powrel & Adhikari, 2022).

For	vdW (kcal/mol)	Electrostatic (kcal/mol)	Sum(kcal/mol)
Equilibration HbS	-1663.02	-7023.26	-8686.28
Equilibration β chain	-440.90	-1519.27	-1960.17

Table 5: The average vdW and electrostatic interaction potential energy in normal hemoglobin protein and its beta chain (Powrel & Adhikari, 2022).

For	vdW (kcal/mol)	Electrostatic (kcal/mol)	Sum (kcal/mol)
Equilibration A3N	-1705.00	-6505.72	-8210.72
Equilibration β chain	- 425.23	-1491.36	-1916.60

The sum of the vdW and electrostatic interaction energy of sickle hemoglobin and normal hemoglobin were found as in the the data of Table 4 and Table 5. When sickle and normal hemoglobin proteins were in equilibrium, the total energy of the both protein and their beta chain have shown higher values in sickle than that of normal hemoglobin (Table 4 and 5). The majority of the vdW and electrostatic interaction energy of both proteins appears to be carried by the beta chain. The tabular figures also show that the electrostatic potential energies of sickle and normal hemoglobin proteins are significantly larger than those of vdW.

vdW and Electrostatic Interaction in Sickle and Normal Hemoglobin Dimer

We have also carried out the quantitative estimation of the non-bonded binding interaction of proteins, β chain (PROB) and α chain (PROA) in the dimers.

Table 6: The averages of vdW and electrostatics interaction potential energy of sickle hemoglobin protein dimer.

For sickle Hb dimer	vdW	Electostat	Total (kcal/mol)
Protein	-884.62	-2719.82	-3604.45
PROA	-393.30	-1239.29	-1632.59
PROB	-457.01	-1443.64	-1900.65
Between PROA and PROB	-32.83	-36.03	-68.86

Table 7: The averages of vdW and electrostatics interaction potential energy of normal hemoglobin protein dimer.

For Normal Hb dimer	vdW	Electostat	Total (kcal/mol)
Protein	-924.05	-2852.37	-3776.41
PROA	-430.92	-1316.29	-1747.21
PROB	-462.27	-1467.60	-1929.87
Between PROA and PROB	-30.93	-67.58	-98.50

We have also estimated the vdW and electrostatic interaction energy in between the chain α and chain β of both the dimers (Table 6 and Table 7) for the confirmation of the structures. The electrostatic interaction energy in between α (PROA) and β (PROB) of the normal hemoglobin dimer was 67.58 kcal/mol and of sickle dimer 36.03 kcal/mol. Due to this, the sum of vdW and electrostatic energy is much higher in normal dimers (98.51 kcal/mol) than in sickle dimers (68.86 kcal/mol).

Our study of vdW and electrostatics interactions in hemoglobin shows that the sickle is more strongly bound than that of normal hemoglobin protein. However, the two chains of dimer in sickle were loosely bounded by vdW and electrostatic energy than that of normal hemoglobin dimer.

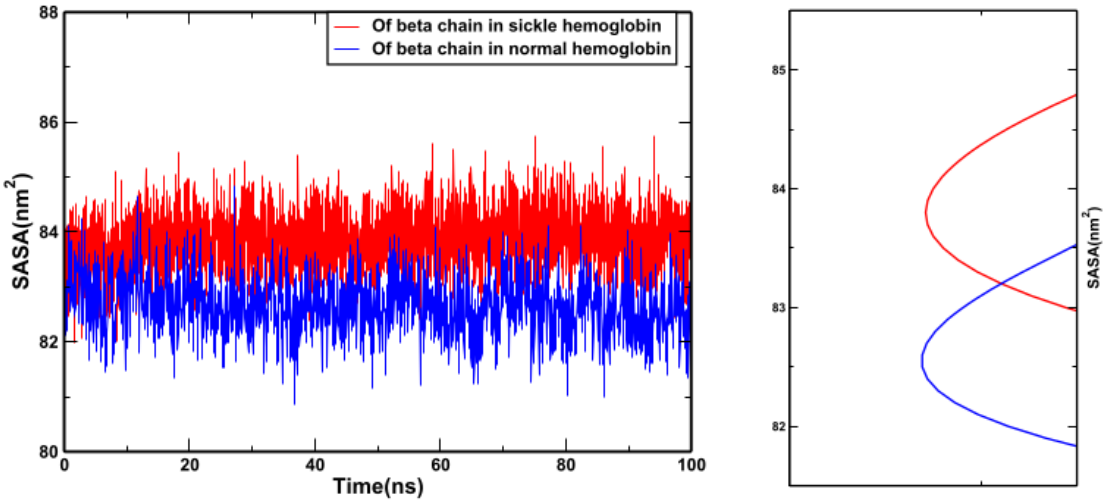
4.2.6 Solvent Accessible Surface Area (SASA) Calculations

The solvent accessible surface area is another component that evaluates the region of the molecular surface exposed enough to the solvent to interact with solvent molecules. The contact surface area of both chains of each system was determined (Zou et al., 2012; Koirala et al., 2021) by Equation 3.33. When analyzing the structural stability of the molecular complex, the contact surface area between two molecules in contact takes a vital role. For this, the average solvent accessible surface area (SASA) of each molecule and its complex was estimated. We have presented the SASA of the beta chain in both sickle and the normal tetramer, chain ACD in sickle and normal tetramer, sickle and normal hemoglobin tetramer, sickle dimer and normal hemoglobin dimer.

SASA During Beta Chain SMD in Tetramer Hemoglobin

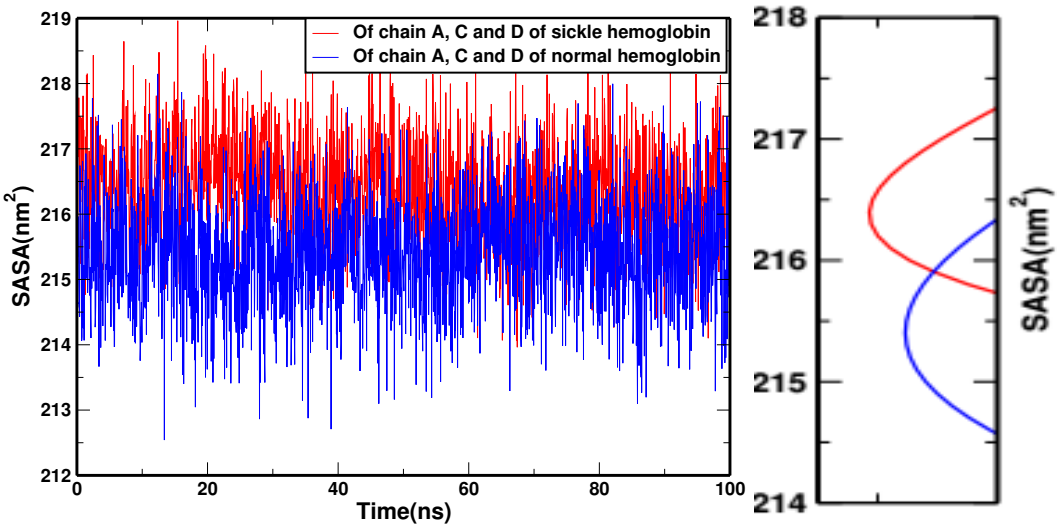
SASA Interaction of Sickle Hb, Normal Hemoglobin and Beta Chain Here, we have estimated the SASA of beta chain in Figure 40, SASA of PROA PROC PROD chain in Figure 41), SASA of sickle and normal hemoglobin protein in Figure 42 and SASA during unfold of protein chains in Figure 43 (a) and (b)).

.



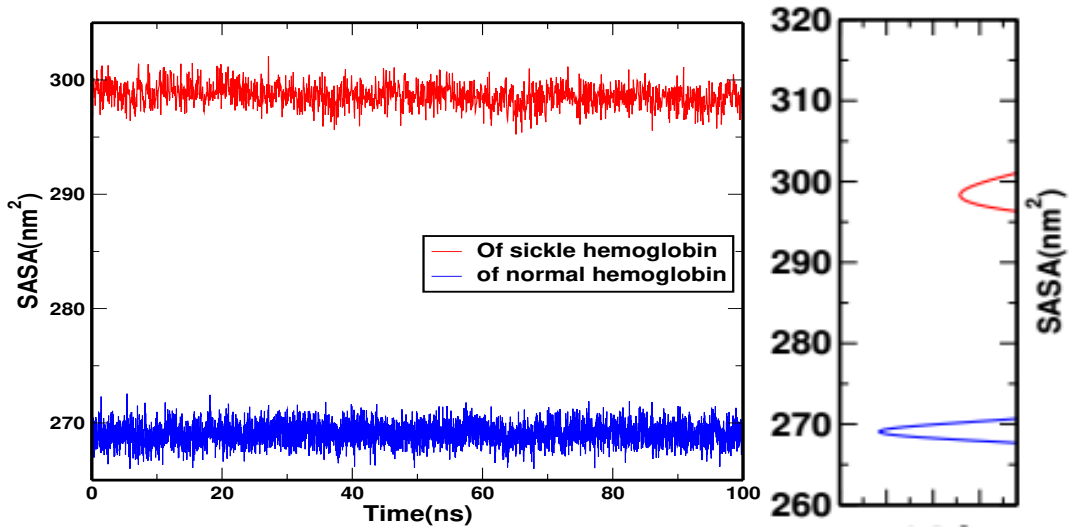
(a) (b)

Figure 40: The SASA (a) of beta chain (chain B) in sickle and normal hemoglobin tetramer and (b) SASA distribution with time (Powrel & Adhikari, 2022).



(a) (b)

Figure 41: SASA (a) of A, C and D chain alone of sickle and normal hemoglobin protein and (b) SASA distribution with time (Powrel & Adhikari, 2022).



(a) (b)

Figure 42: SASA (a) of sickle and normal hemoglobin protein tetramer and (b) SASA distribution.

Figure 41 and 42 are the SASA plots of ACD chains and the protein chain of both hemoglobin tetramer system. In both figures it can be seen that the subunits of sickle protein have a higher surface area in contact with water (Powrel & Adhikari, 2022). i.e., the sickle is less hydrophobic than the normal. The figure 42 (a) and (b) are the SASA plots of whole sickle and normal hemoglobin protein and their SASA distribution respectively. The plot confirms the total SASA of hemoglobin protein is the sum of the SASA's of their sub units. The Figure 42 indicates the peak value of the SASA distribution of sickle and normal Hb tetramer lies at 299 nm^2 and 269 nm^2 respectively.

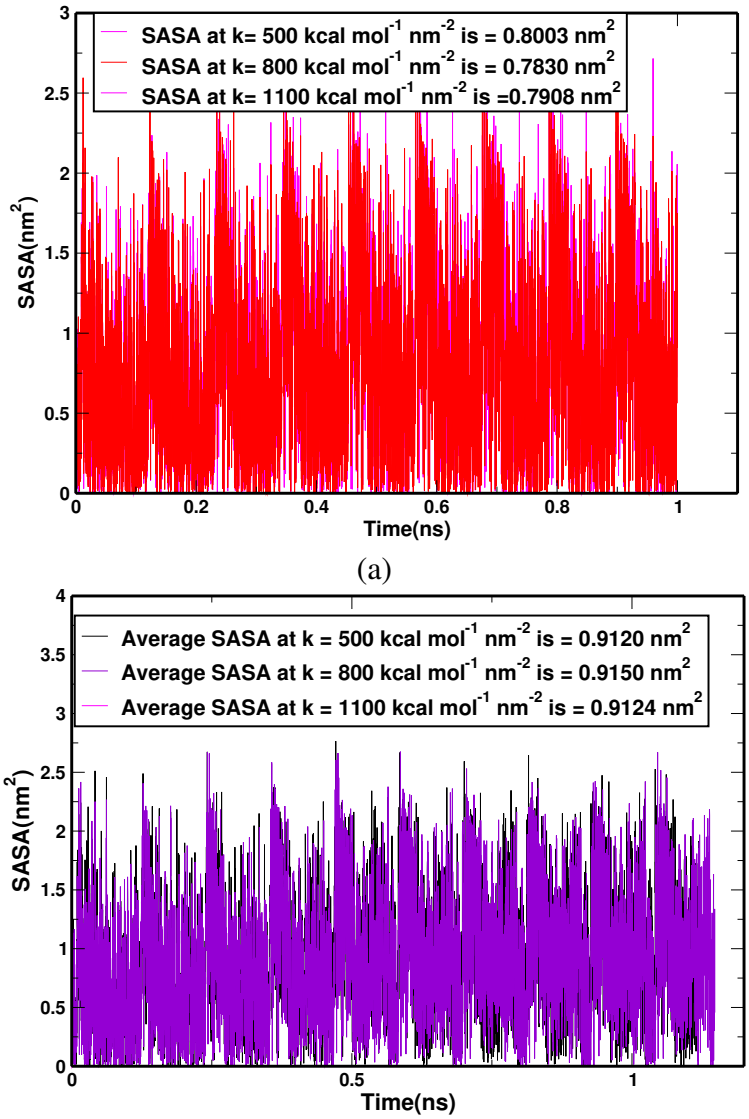


Figure 43: (a) Plot of SASA of equilibrated beta chain in sickle protein complex (b) Plot of SASA of equilibrated beta chain in normal hemoglobin protein complex for for k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022).

Table 8: SASA of HbS and A3N protein with k=5, k=8 and k=11 kcal mol $^{-1}$ Å^{-2}.

SASA	SASA(nm^2)			Net contact area (S) nm^2
of	Chain B (S_1)	Chain ACD(S_2)	Protein(S_3)	= (S_1+S_2-S_3)/2
HbS	83.8514	215.3401	269.0876	15.0520
A3N	82.6633	216.4070	298.6180	0.2261

SASA is increasing with the time scale for different values of spring constants (k) and becomes constant (Figure 43 (a) and (b)). At equilibrium, SASA of sickle hemoglobin protein (Table 8) is 269.08 nm^2 and contact area of interface of PRO B and PRO ACD is 15.0520 nm^2. During SMD, the SASA of β chain of HbS is decreased to around 0.80 nm^2 (Figure 43 (a)). In normal hemoglobin, at equilibrium the SASA is

298.61 nm^2 (Table 8) and is decreased during SMD of β chain in normal hemoglobin to around 0.90 nm^2 (Figure 43 (b). The contact area of interface of PRO B and PRO ACD in normal Hb is found to be is 0.23 nm^2. Higher the k values the SASA is also higher which indicates that the beta chain of sickle hemoglobin protein is hydrophilic in nature. There was much difference in SASA of normal and sickle hemoglobin (Figure 42 (a) and 42 (b)). It indicates the chances of a sharp reduction in inter-facial tension (IFT) of normal hemoglobin than of sickle hemoglobin (Andersson et al., 2014).

SASA Interaction of Sickle Hb, Normal Hb Protein and Alpha Chain

SASA calculation is essential indicator for finding the contact area of two individual constituents and stability of the relaxed protein (Möller et al., 2017) of the molecules. We have also estimated the SASA of sickle hemoglobin protein, normal hemoglobin protein, chain A (PROA) of both protein and of chain BCD (PROB PROC PROD) of tetramer structure separately. An average value of contact surface area of constituent molecules and their complex were determined

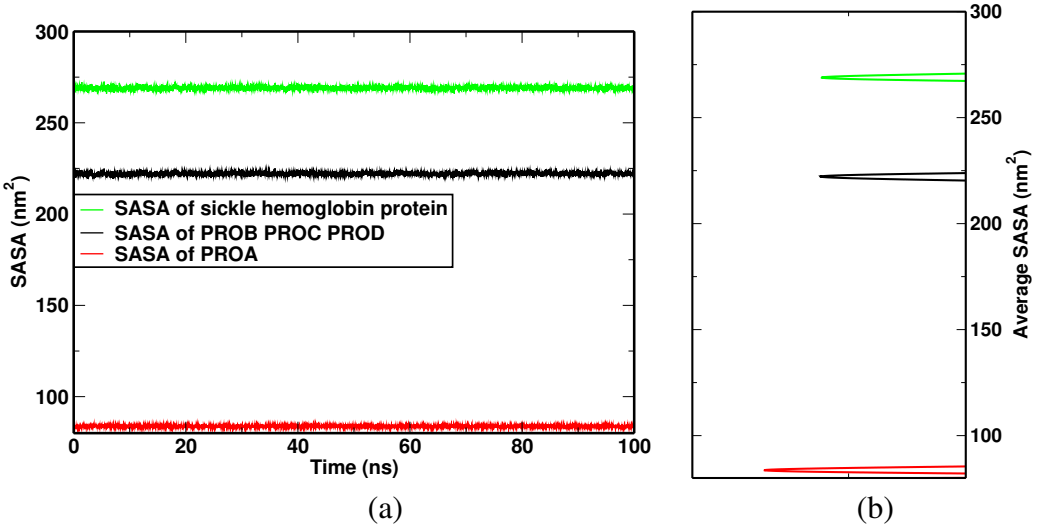


Figure 44: SASA of a) sickle HbS protein and b) SASA distribution with time..

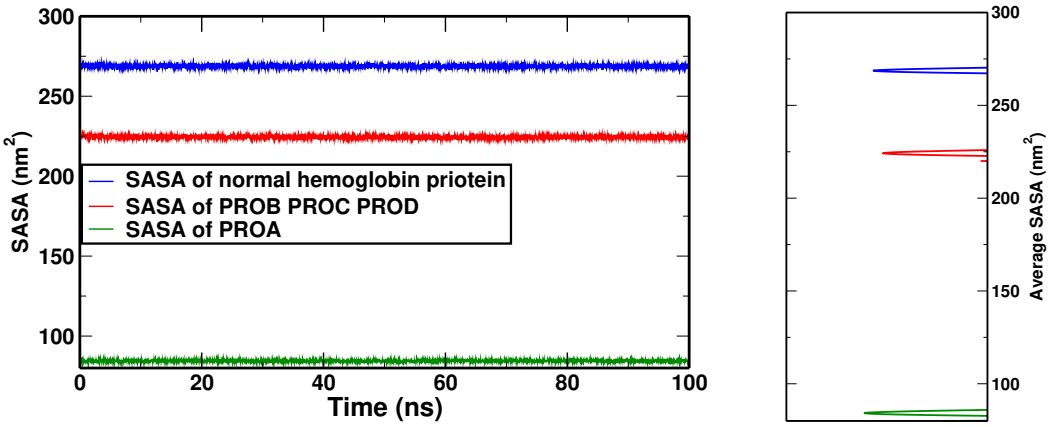


Figure 45: SASA of a) normal Hb protein and b) SASA distribution with time.

Table 9: SASA distribution with time of chain α (A), chain BCD and protein of both sickle and normal hemoglobin dimer in nm^2.

SASA	Chain A	Chain BCD	Hb protein	Contact Area=($S_1 + S_2 - S_3$)/2
of	S_1 (nm) ^{2}	S_2(nm) ^{2}	S_3 (nm) ^{2}	S(nm) ^{2}
Sickle	83.7539	222.1547	269.0876	18.4105
Normal	84.4384	224.4335	268.7659	20.0530

The SASA estimation of sickle hemoglobin protein are in Figure 44 (a) and Figure 45 (a). Figure 44 (b) and 45 (b) are the average SASA of sickle and normal hemoglobin tetramer for the system of alpha SMD. The total SASA for the sickle protein is $\approx$ 269.08 nm ^{2} (Table 9). The SASA of chain A and chain BCD lies nearly at $\approx$ 83.75 nm ^{2} and 222.15 nm ^{2} respectively.

SASA of normal hemoglobin protein is also estimated (Figure 45 (a)). The average of the SASA are shown in Figure 45 (b). The total SASA of the normal hemoglobin protein found to be $\approx$ 268.76 nm ^{2} which is slightly less than of sickle (Table 9). It is also seen that the SASA of chain A and chain BCD in normal hemoglobin protein are much higher than of sickle protein and lies around 84.43 nm ^{2} and 224.43 nm ^{2} respectively. The high reduction of contact area in between chain A and chain BCD of sickle hemoglobin indicating the release of energy difference of binding as free energy.

SASA of Sickle and Normal Hemoglobin Dimer

We have also estimated the SASA in the dimer structure of both sickle and normal hemoglobin protein. The SASA is further useful in estimating the contact surface area of two molecules. Contact surface area of two chains of a molecule is the surface at inter-facial region that contributes in binding two chains. Greater is the contact surface area, greater is the binding strength between the chains. This assists to find the stability of complex formation of bio-molecule.

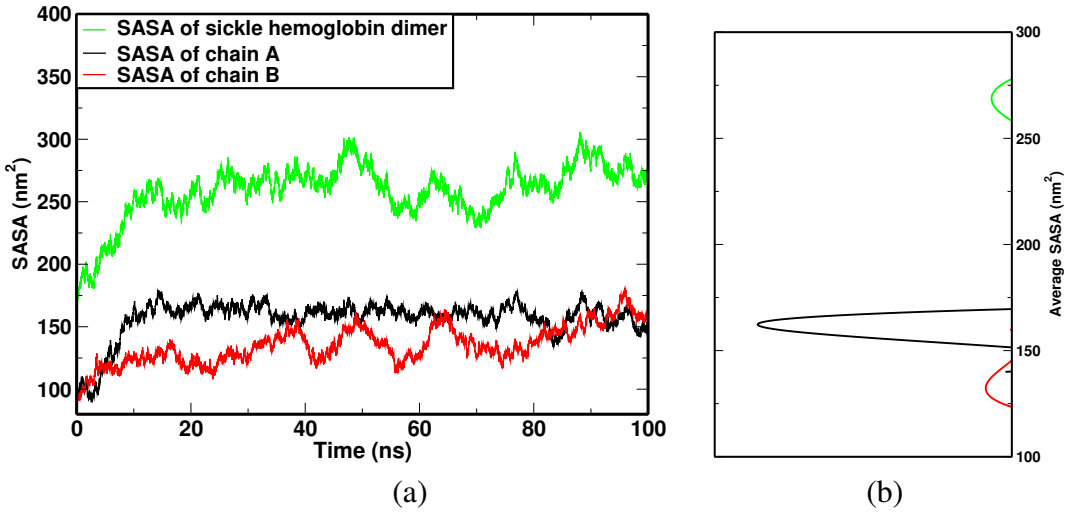


Figure 46: (a) SASA's in sickle hemoglobin dimer and (b) SASA distribution with time (Powrel et al., 2022)

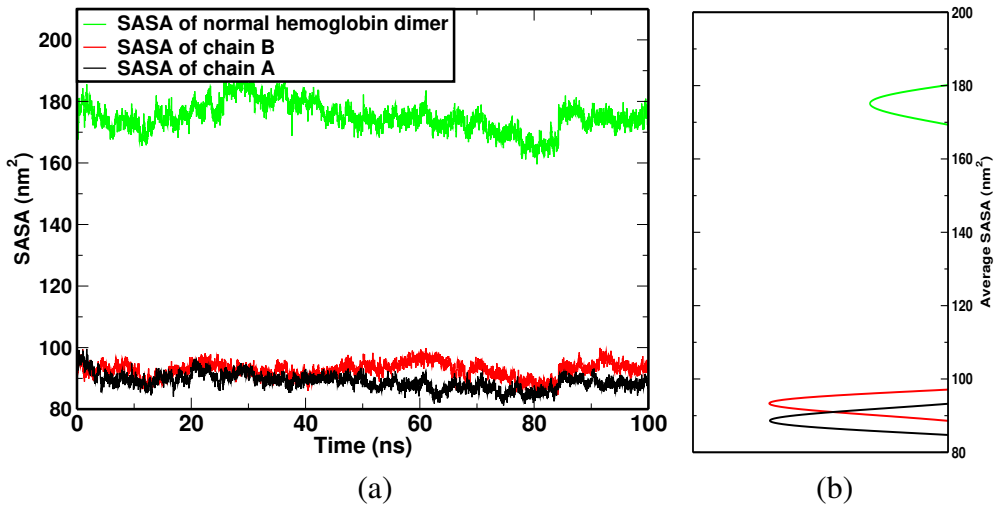


Figure 47: (a) SASA's in normal hemoglobin dimer and (b) SASA distribution with time (Powrel et al., 2022).

Table 10: SASA distribution with time of α chain, β chain andwhole proteins of both sickle and normal hemoglobin dimer in Å^2 (Powrel et al., 2022).

SASA	Chain A	Chain B	Hb protein	Contact Area (nm) ^{2}
of	S_1 (nm) ^{2}	S_2 (nm) ^{2}	S_3 (nm) ^{2}	$=(S_1 + S_2 - S_3)/2$
Sickle dimer	154.3312	138.2752	259.7974	16.4045
Normal dimer	86.3563	92.7481	174.9669	2.0688

It indicates that the SASA of sickle protein is found to be 259.79 nm^2 and of normal hemoglobin is at 174.96 Å^2 (Table 10). Higher SASA means higher is the contact area of two chains in the dimer. The high reduction of contact area in between chain A and chain B of normal hemoglobin dimer indicating the release of energy difference of binding as free energy.

4.2.7 Elastic Property by Steered Molecular Dynamics

The mechanical functions of protein can be studied by SMD techniques. The steered molecular dynamics (SMD) study of protein has been performed for both sickle and normal hemoglobins for estimating the force required to break the hydrogen bonding of α and β chains in the hemoglobin molecules. Such analysis provides the information about the contribution of α and β chain to the elastic property of the hemoglobin protein molecules. The present work focuses on the estimation of the forces required in breaking the hydrogen bond to understand the elastic stiffness in sickle and normal hemoglobin structures using steered molecular dynamics simulation. This SMD is the study of the elastic stiffness in the sickle and normal hemoglobin protein due to the presence of hydrogen bonds in the hemoglobin chains.

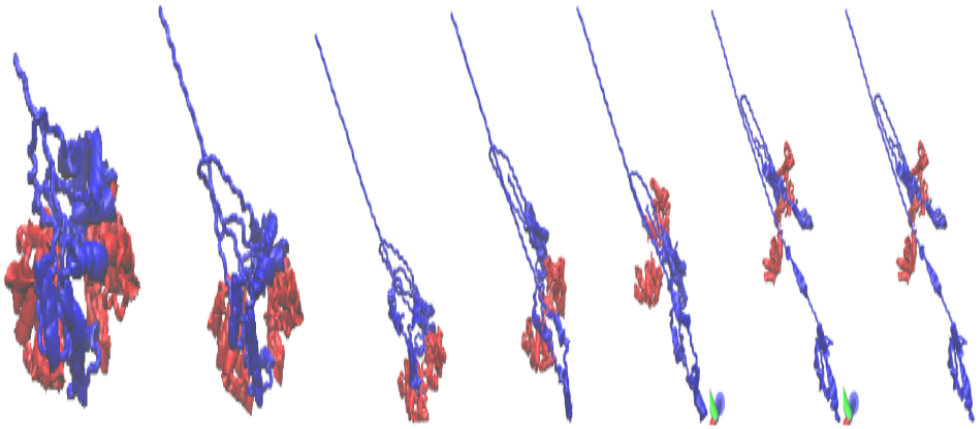


Figure 48: Snapshots of SMD of alpha chain of sickle hemoglobin protein.

After studying the SMD of the protein chain, it looks essential to estimate the presence of H- bonding in the protein chains to conclude on the binding force in both proteins. Set of Figure 48 is the seven consecutive snapshots taken during NC terminal SMD of hemoglobin protein.

SMD of Beta Chain of Sickle and Normal Hemoglobin Protein

N-C terminal SMD study of β chain provides the elastic property of the hemoglobin protein due to β chain alone. It can explore unfolding paths of the proteins and show the contribution of the β chain to the stiffness of the whole Hb complex. In this N-C terminal β chain SMD, we have pulled the one terminal of the beta chain by keeping the other end fixed. In the Figure 49, we have shown the beta SMD graph of force required versus time with spring constants of5 kcal mol^{-1} Å^{-2} in figures 49 (a), (b) and (c) respectively for sickle hemoglobin and figures 49 (d), (e) and (f) respectively for normal hemoglobin (Powrel & Adhikari, 2022).

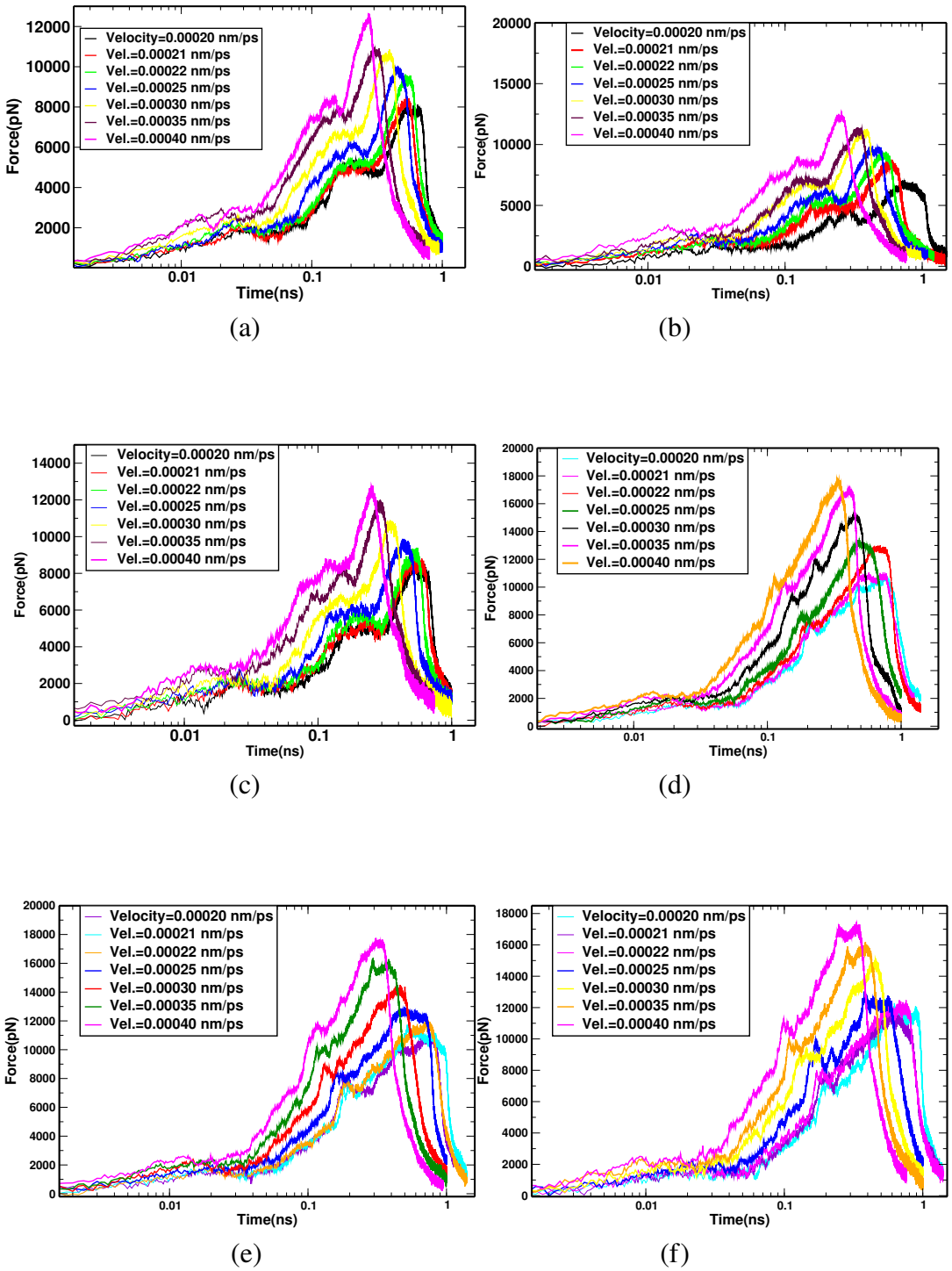


Figure 49: SMD graph of the beta chain in HbS protein with a) k=5, b) k=8, c) k=11 kcal mol $^{-1}$ Å^{-2} and d) SMD graph of A3N protein with k=5 kcal mol $^{-1}$ Å^{-2} e) SMD graph of A3N protein with k=8 and f) SMD graph of A3N protein with k= 11 kcal mol $^{-1}$ Å^{-2} (Powrel & Adhikari, 2022).

The peak of the SMD plots indicate the maximum force required for breaking the bindings of the molecules in the complex. From the plots, it is seen that the values of peak forces increases with the increase in pulling velocity as well as spring constant. Higher pulling velocity results in the loss of equilibrium states and significant error in simulation. The SMD graph of sickle hemoglobin at a spring constant of k=5 kcal mol $^{-1}$ Å^{-2}

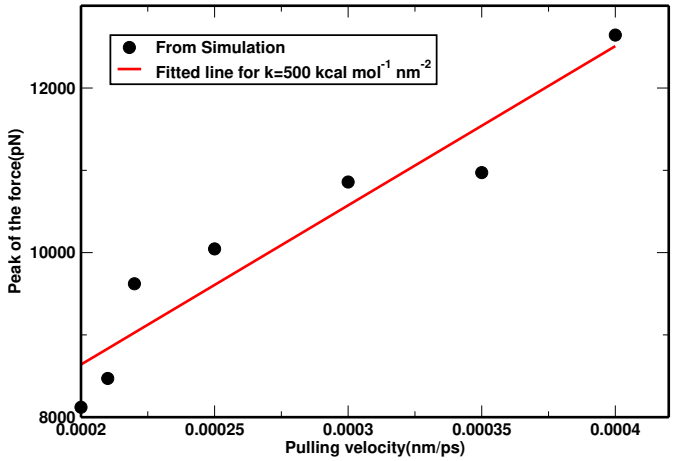
indicates that at much lower pulling velocity, the peak is flat with the least value. As pulling velocity is increased, it is gradually increased i.e. stiffness is high at low k. As spring constant is increased from k=5 kcal mol $^{-1}$ Å^{-2} to k=8 kcal mol $^{-1}$ Å^{-2} and k=11 kcal mol $^{-1}$ Å^{-2}, the corresponding forces required for breaking H-bond are also increased. Similarly, for the increased spring constant from k=5 kcal mol $^{-1}$ Å^{-2} to k=8 kcal mol $^{-1}$ Å^{-2} and k= 11 kcal mol $^{-1}$ Å^{-2}, the corresponding force required for breaking H-bond is also increased in normal hemoglobin. Again, the rupture force fluctuated greatly when spring constant was larger than 11 kcal mol $^{-1}$ Å^{-2}. Meanwhile, the effect was not obvious when spring constants were less than 5 kcal mol $^{-1}$ Å^{-2}. The peak of the forces required are tabulated in Table 11.

Table 11: Forces (pN) for H-bond breaking in sickle hemoglobin (HbS) and normal hemoglobin (A3N) with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022).

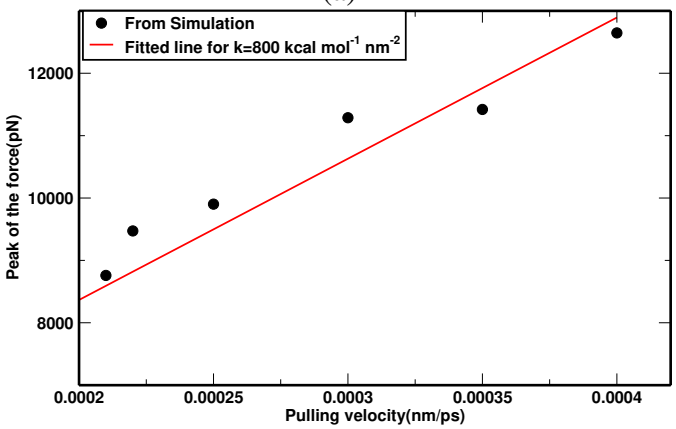
Force required for	HbS (with k in kcal mol^{-1} Å^{-2})			A3N (with k in kcal mol^{-1} Å^{-2})		
Velocity (Å/ps)	k=5	k=8	k=11	k=5	k=8	k=11
0.0020	8120.70	7073.74	8908.22	10809.90	12005.00	10425.40
0.0021	8470.74	8759.14	8984.77	10881.50	12577.10	12820.80
0.0022	9621.09	9471.70	9397.36	12421.30	11761.80	11336.90
0.0025	10045.30	9902.31	9881.79	13396.80	12951.60	13032.60
0.0030	10858.20	11286.50	10900.00	15276.50	14473.10	15140.90
0.0035	10972.00	11420.10	12015.00	17113.30	16359.40	16187.20
0.0040	12643.40	12646.80	12770.50	17676.70	17753.70	17434.10

From Table 11, it is seen that for a particular spring constant, as the pulling velocity is increased the peak of the forces are also increased. In addition, the peak of the forces are also increased for increased in the values of spring constants. In order to check the nature of variation of peak force with the pulling velocity. For k=5 kcal mol^{-1} Å^{-2}, while plotting the peak of force versus the corresponding pulling velocity a straight line is obtained as in Figure 50 (a), for k=8 kcal mol^{-1} Å^{-2}, the peak of the force versus the pulling velocity gives another straight line as in Figure 50 (b) and for k=11 kcal mol^{-1} Å^{-2}, the peak of the force versus the pulling velocity gives a straight line as in Figure 50 (c). The slope of these three straight line are compared in Figure 50 (a).

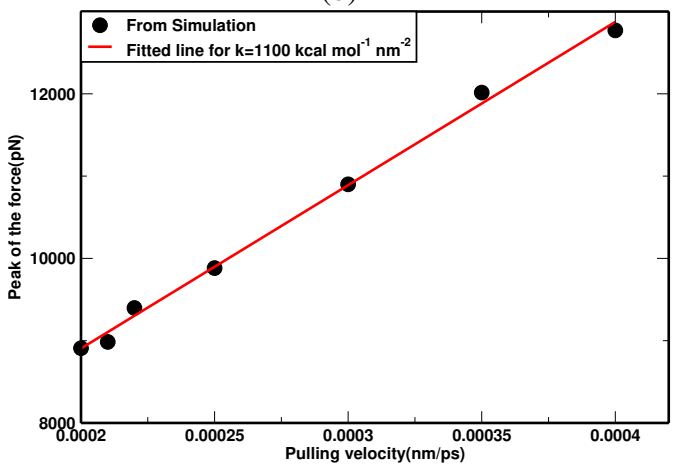
Again for the normal hemoglobin we have plotted the peak of the forces versus the pulling velocity with spring constant of k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} as shown in Figures 51 (a), (b) and (c) respectively (Powrel & Adhikari, 2022).



(a)

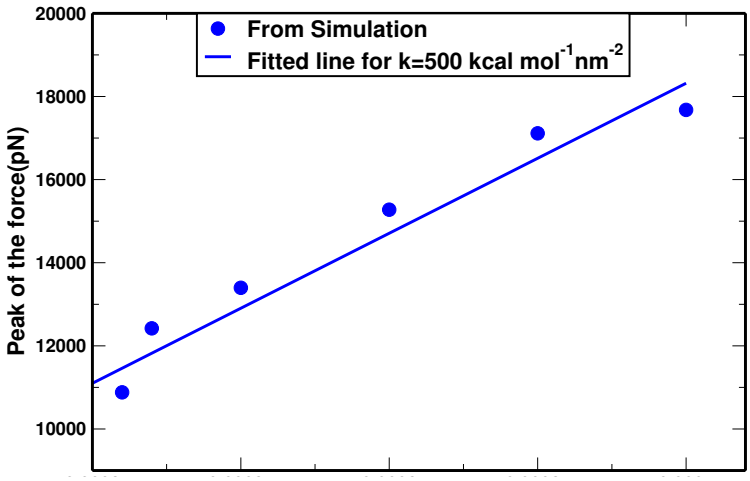


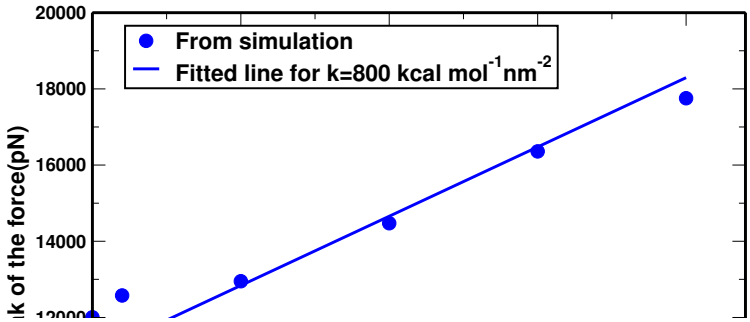
(b)

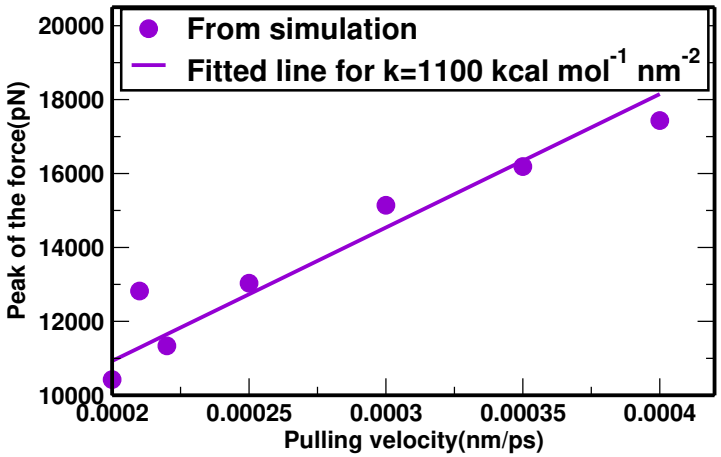


(c)

Figure 50: a) Pulling velocity versus force required graph of HbS with k=5 kcal mol^{-1} Å^{-2} b) Pulling velocity versus force required graph of HbS with k=8 kcal mol^{-1} Å^{-2} kcal mol^{-1} Å^{-2} c) Pulling velocity versus force required graph of HbS with k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022).







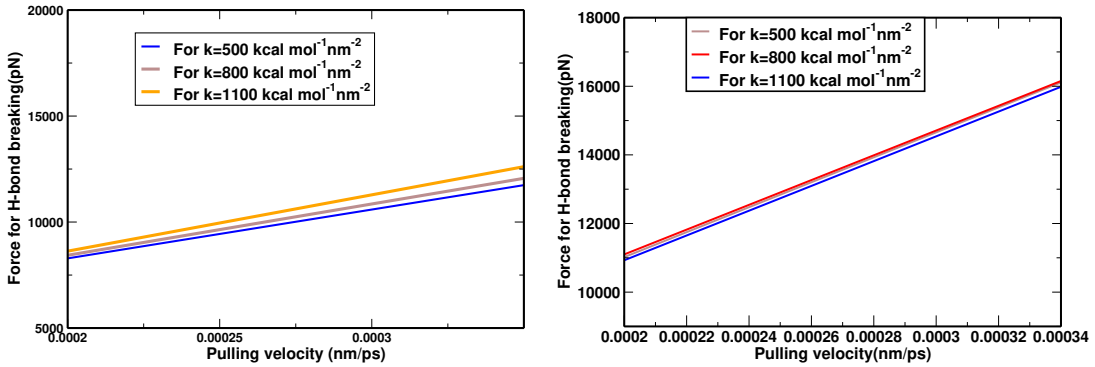
(c)

Figure 51: a) Pulling velocity versus force required graph of A3N with k=5 kcal mol^{-1} Å^{-2} b) Pulling velocity versus force required graph of HbS with k=8 kcal mol^{-1} Å^{-2} and c) Pulling velocity versus force required graph of HbS with k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022).

The effects of pulling velocity on forces required in breaking the hydrogen bonding

of sickle hemoglobin protein with different spring constant k are shown (Table 11) in Figure 50 (a) and (b) and (c). They showed that the rupture force had obvious linear correlations when the pulling velocity ranged from 0.002 Å/ps to 0.004 Å/ps. Therefore, a minimum pulling velocity of 0.002 Å/ps is chosen in the SMD simulation of HbS. The rupture force is also affected by the pulling velocity at various spring constant k (Table 11) for normal hemoglobin as shown in Figure 51 (a) and (b) and (c). They shows the linear correlation when the pulling velocity ranged from 0.002 Å/ps to 0.004 Å/ps in normal hemoglobin. Therefore minimum pulling velocity was taken as 0.002 Å/ps in the case of normal hemoglobin too. As spring constant k is increased the stiffness is more in A3N for lower k.

To understand the variation in slope of the straight lines of the set of Figures 50 (a), (b) and (c) of sickle hemoglobin, we have compared them in Figure 52 (a). Similarly, we have compared the slope of the straight lines of Figure 51 (a), (b) and (c) of normal hemoglobin as in Figure 52 (b).



(a) (b)

Figure 52: Pulling velocity versus force required graph for beta chain of (a) HbS with different k with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2}.and (b) A3N with different with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} (Powrel & Adhikari, 2022).

On plotting the graph between pulling velocity and peak of the force needed for breaking the hydrogen bond from the values in Table 11; Figure 52 (a) and (b) are generated for HbS and *A3N* respectively. For sickle hemoglobin with different values of k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2}, the pulling velocity versus force required graph possesses the higher slope as the pulling velocity is increased, i.e. the variation of force occurs with uniform rise in the velocity and spring constant (Figure 52 (a). It also indicates that more the value of spring constant more is the force required i.e. stiffness is higher. Instead, for different values of k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} of A3N, the velocity versus force graph possess the same slope. It shows that the force varies with uniform variation in the velocity and spring constant (Figure 52 (b).

The above study reveals that the hydrogen bond interaction force of the beta chain is

7073.74 pN to 12646.80 pN for pulling velocities of 0.00020 nm/ps to 0.00040 nm/ps with the spring constant of 800 kcal mol^{-1} nm^{-2} in the case of the sickle hemoglobin protein. However, in normal hemoglobin protein, the hydrogen bond interaction force in the beta chain is ranging from 12005.00 pN to 17753.70 pN for pulling velocities of 0.00020 nm/ps to 0.00040 nm/ps using the spring constant of 800 kcal mol^{-1} nm^{-2}.

With different k; (k=500 kcal mol^{-1} nm^{-2}, k=800 kcal mol^{-1} nm^{-2} and k=1100 kcal mol^{-1} nm^{-2}) the velocity versus force graph of sickle hemoglobin shows the increased slope for increasing velocity. It means the required force for breaking the hydrogen bond varies with uniform change in the velocity and spring constant (Figure 52 (a)). It also indicates that more the value of the spring constant, more is the force required i.e. stiffness is higher. The velocity versus force graph with different values of k: (k=500 kcal mol^{-1} nm^{-2}, k=800 kcal mol^{-1} nm^{-2} and k=1100 kcal mol^{-1} nm^{-2}) for A3N are also plotted. They have got the same slope for each of the various k. This indicated the variation of force with uniform change in velocity and spring constant (Figure 52 (b)).

SMD study of Alpha Chain in Sickle and Normal Hemoglobin Protein

N-C terminal SMD study of alpha chain also provides the elastic property of the hemoglobin protein due to alpha chain alone. They can explore unfolding paths of the proteins and show the contribution of the α chain to the stiffness of the whole Hb complex. In this N-C terminal α chain SMD also, we have pulled the one terminal of the alpha chain by keeping the other end fixed. In the Figures 53 (a), (b) and (c), we have shown the alpha SMD graph of force required versus time with spring constants of k=5, 8 and 11 kcal mol^{-1} $Å^{-2}$ respectively for sickle hemoglobin and Figures 53 (d), (e) and (f) for normal hemoglobin.

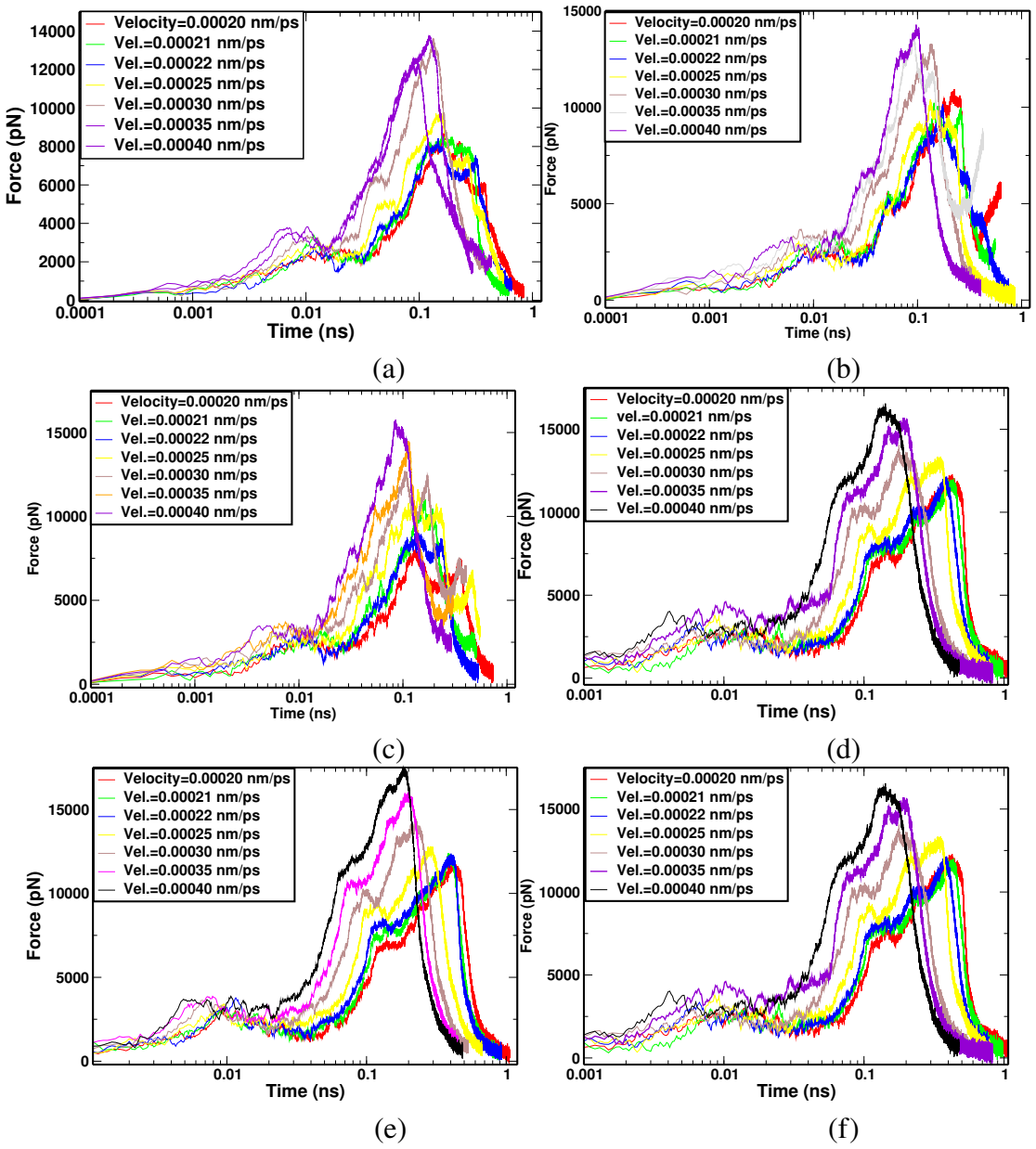


Figure 53: SMD graph of alpha chain in HbS protein with a) k=5, b) k=8, c) k= 11 kcal mol $^{-1}$ Å^{-2}, d) SMD graph of alpha chain in A3N protein with k=5 kcal mol $^{-1}$ Å^{-2}, e) SMD graph of A3N protein with k=8 and f) SMD graph of A3N protein with k= 11 kcal mol $^{-1}$ Å^{-2}.

The peak of the SMD plots also indicate the maximum force required for breaking the bindings of the molecules in the complex. From the plots, it is seen that the values of peak forces increase with the increase in pulling velocity as well as spring constant. Higher pulling velocity results in the loss of equilibrium states and significant error in simulation. The alpha SMD graph of sickle hemoglobin at a spring constant of k=5 kcal mol $^{-1}$ Å^{-2} indicates that at much lower pulling velocity, the peak is flat with the least value. As pulling velocity is increased, the peak of the force is gradually increased i.e. stiffness is found high at low k in alpha SMD. As spring constant is increased from k=5 kcal mol $^{-1}$ Å^{-2} to k=8 kcal mol $^{-1}$ Å^{-2} and k=11 kcal mol $^{-1}$ Å^{-2}, the corresponding

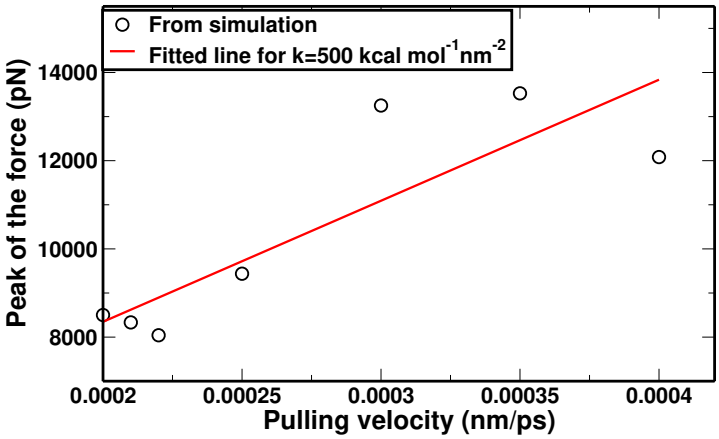
forces required for breaking H-bond are also increased in sickle hemoglobin. Similarly, for the increased spring constant from k=5 kcal mol $^{-1}$ Å^{-2} to k=8 kcal mol $^{-1}$ Å^{-2} and k= kcal mol $^{-1}$ Å^{-2}, the corresponding force required for breaking H-bond is also increased in normal hemoglobin. Again, the rupture force fluctuated greatly when spring constant was larger than 11 kcal mol $^{-1}$ Å^{-2}. Meanwhile, the effect was not obvious when spring constants were less than 5 kcal mol $^{-1}$ Å^{-2}. The peak of the forces required are tabulated in Table 12.

Table 12: Forces (pN) required for breaking the H-bonds in the alpha chain of sickle (HbS) and normal hemoglobin (A3N) protein with k=5 kcal mol^{-1} Å^{-2}, k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2}.

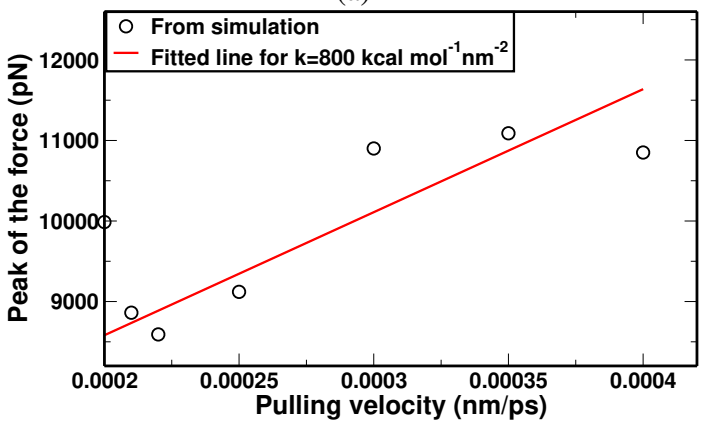
Force (pN)	HbS (with k in kcal mol^{-1} Å^{-2})			A3N (with k in kcal mol^{-1} Å^{-2})		
Velocity (nm/ps)	k=5	k=8	k=11	k=5	k=8	k=11
0.00020	8500.40	9986.65	6735.06	11288.40	11343.20	11608.80
0.00021	8332.78	8861.26	9645.39	11528.60	11536.90	10939.60
0.00022	8041.71	8591.47	8080.73	11543.00	11547.20	11024.50
0.00025	9436.98	9119.81	10021.20	11929.40	11845.30	13349.70
0.00030	13250.80	10900.64	10379.40	13396.00	13188.70	12560.00
0.00035	13527.50	11089.51	10807.40	14258.90	14426.40	14243.70
0.00040	12082.01	10850.87	12345.50	15114.10	15319.60	16557.10

From Table 12, it is seen that for a particular spring constant, as the pulling velocity is increased the peak of the forces are also increased. In addition, the peak of the forces are also increased for increased in the values of spring constants. For k=5 kcal mol^{-1} Å^{-2}, while plotting the peak of force versus the corresponding pulling velocity in sickle hemoglobin a straight line is obtained as in Figure 54 (a), for k=8 kcal mol^{-1} Å^{-2}, the peak of the force versus the pulling velocity gives another straight line as in Figure 54 (b) and for k=11 kcal mol^{-1} Å^{-2}, the peak of the force versus the pulling velocity gives a straight line as in Figure 54 (c). The slope of these three straight lines is compared in Figure 56 (a).

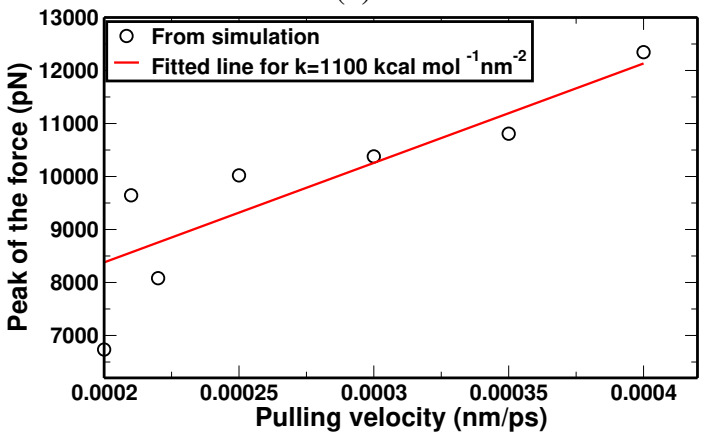
Again for the normal hemoglobin we have plotted the peak of the forces versus the pulling velocity with spring constant of k=5 kcal mol^{-1} Å^{-2} , k=8 kcal mol^{-1} Å^{-2} and k=11 kcal mol^{-1} Å^{-2} as shown in Figures 55 (a), (b) and (c) respectively.



(a)

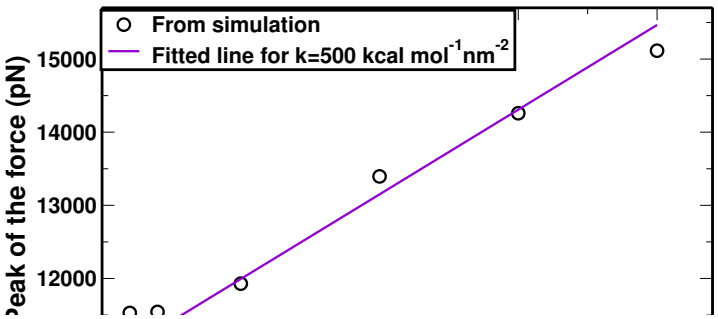


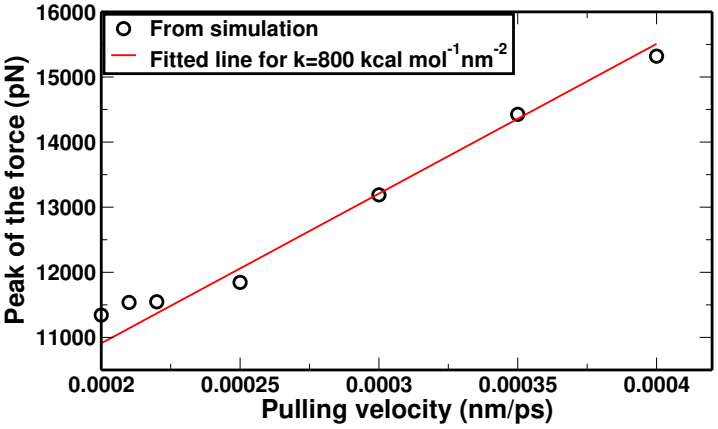
(b)



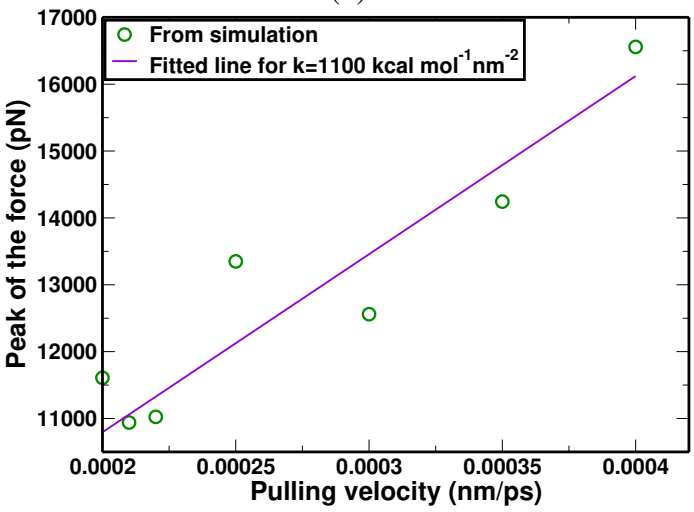
(c)

Figure 54: (a) Pulling velocity versus force required graph of HbS with K=500 (b) Pulling velocity versus force required graph of HbS with K=800 kcal $mol^{-1}nm^{-2}$ and (c) Pulling velocity versus force required graph of HbS with K=1100 kcal $mol^{-1}nm^{-2}$.





(b)



(c)

Figure 55: (a) Pulling velocity versus force required graph of A3N with k=500 (b) Pulling velocity versus force required graph of A3N with k=800 kcal $mol^{-1}nm^{-2}$ and (c) Pulling velocity versus force required graph of A3N with k=1100 kcal $mol^{-1}nm^{-2}$.

The effects of pulling velocity on forces required in breaking the hydrogen bonding of sickle hemoglobin protein with different spring constant k are shown (Table 12) in Figure 54 (a) and (b) and (c). They showed that the rupture force had obvious linear correlations when the pulling velocity ranged from 0.002 Å/ps to 0.004 Å/ps. Therefore,

a minimum pulling velocity of 0.002 Å/ps is chosen in the SMD simulation of HbS.

The rupture force is also affected by the pulling velocity at various spring constant k (Table 12) for normal hemoglobin as shown in Figure 55 (a) and (b) and (c). They shows the linear correlation when the pulling velocity ranged from 0.002 Å/ps to 0.004 Å/ps in normal hemoglobin. Therefore minimum pulling velocity was taken as 0.002 Å/ps in the case of normal hemoglobin too. As spring constant k is increased the stiffness is more in A3N for lower k.

To compare the variation in slope of the straight lines of the set of Figures 53 (a), (b) and (c) of sickle hemoglobin, we have plotted them in Figure 56 (a). Similarly, we have compared the slope of the straight lines of Figure 55 (a), (b) and (c) of normal hemoglobin as in Figure 56 (b).

The Figure 50 (a), (b) and (c) show the impact of pulling velocity on the force required for various spring constant k in sickle hemoglobin. They showed that the rupture force had a linear correlation when the pulling velocity ranged from 0.0002 nm/ps to 0.0004 nm/ps. It seems that the minimum pulling velocity for the SMD simulation is 0.0002 nm/ps for sickle hemoglobin protein. Similarly, the rupture force for the different pulling velocities at various spring constant k for normal hemoglobin (Table 12) is in Figure 53 (a), (b) and (c). They also obvious the linear correlations with the pulling velocity ranged from 0.0002 nm/ps to 0.0004 nm/ps in normal hemoglobin. Therefore, the minimum pulling velocity was considered as 0.0002 nm/ps in both the cases of sickle and normal hemoglobin. As spring constant k is increased the stiffness is more in A3N for lower k. We have plotted the graph of the average of the peak of the force needed for breaking the hydrogen bond and pulling velocity with the values from Table 12 which gives Figure 56 (a) and Figure 56 (b) for HbS and A3N respectively.

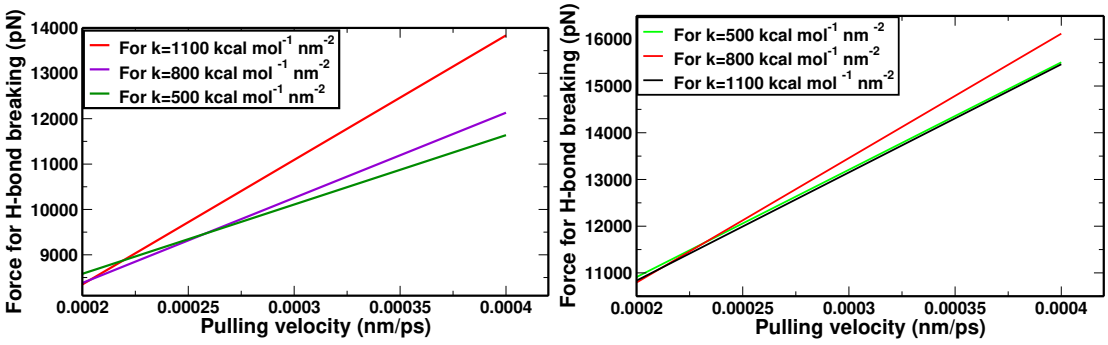


Figure 56: Pulling velocity versus force required graph for alpha chain SMD at different k of k=500 kcal mol^{-1} nm^{-2}, k=800 kcal mol^{-1} nm^{-2} and k=1100 kcal mol^{-1} nm^{-2} for (a) HbS protein and (b) A3N protein.

On plotting the graph between pulling velocity and peak of the force needed for breaking

the hydrogen bond from the values in Table 12; Figure 56 (a) and (b) are generated for HbS and A3N respectively. For sickle hemoglobin with different values of k=5 kcal mol^{-1} $Å^{-2}$, k=8 kcal mol^{-1} $Å^{-2}$ and k=11 kcal mol^{-1} $Å^{-2}$, the pulling velocity versus force required graph possesses the higher slope as the pulling velocity is increased, i.e. the variation of force occurs with uniform rise in the velocity and spring constant (Figure 56 (a). It also indicates that the more the value of the spring constant more is the force required i.e. stiffness is higher. Instead, for different values of k=5 kcal mol^{-1} $Å^{-2}$, k=8 kcal mol^{-1} $Å^{-2}$ and k=11 kcal mol^{-1} $Å^{-2}$ of A3N, the velocity versus force graph possesses the same slope. It shows that the force varies with uniform variation in the velocity and spring constant (Figure 56 (b).

The above study reveals that the hydrogen bond interaction force of the beta chain is 7073.74 pN to 12646.80 pN (Powrel & Adhikari, 2022) for pulling velocities of 0.00020 nm/ps to 0.00040 nm/ps with the spring constant of 800 kcal mol^{-1} nm^{-2} in the case of the sickle hemoglobin protein. However, in normal hemoglobin protein, the hydrogen bond interaction force in the beta chain is ranging from 12005.00 pN to 17753.70 pN for pulling velocities of 0.00020 nm/ps to 0.00040 nm/ps using the spring constant of 800 kcal mol^{-1} nm^{-2}.

With different k; (k=500 kcal mol^{-1} nm^{-2}, k=800 kcal mol^{-1} nm^{-2} and k=1100 kcal mol^{-1} nm^{-2}) the velocity versus force graph of sickle hemoglobin shows the increased slope for increasing velocity. It means the required force for breaking the hydrogen bond varies with uniform change in the velocity and spring constant (Figure 56 (a)). It also indicates that more the value of spring constant more is the force required i.e. stiffness is higher. The velocity versus force graph with different values of k (k=500 kcal mol^{-1} nm^{-2}, k=800 kcal mol^{-1} nm^{-2} and k=1100 kcal mol^{-1} nm^{-2}) for $A3N$ are also plotted. They have got the same slope for each of the various k. This indicated the variation of force with uniform change in velocity and spring constant (Figure 56 (b)).

On comparing the rupture forces for the pulling velocity of 0.00030 nm/ps and spring constant of 800 kcal mol^{-1} nm^{-2}, we got certain conclusions. The SMD of alpha chain (α_1) in tetramer structure of sickle Hb protein indicating that a peak force of 10900.64 pN (Table 12) is required to rupture the H-bonds of intra-chain as well as of inter-chains (among the other three of β_1, α_2, and β_2). Unlike the alpha SMD of normal hemoglobin showing the rupture force of 13188.70 pN (Table 12) for breaking the same intra chain and inter-chain H-bonding. It indicates that the alpha chain is stiffer in normal hemoglobin.

Again,with the same pulling velocity and spring constant, the SMD of beta chain in tetramer structure estimates the rupture force of 12286.50 pN for sickle hemoglobin. Whereas in normal hemoglobin it was found to be 14473.10 pN. i.e., in both SMD study

of alpha and beta chain it is seen that normal hemoglobin has require higher pulling force for breaking the hydrogen bonds. It is also seen that, in beta SMD of tetramer Hb, hydrophobic interaction guides the transcription factors for increase the hydrogen bonding. It is confirmed as per increase in the rupture forces required in beta SMD. However in, sickle Hb, it is comparatively less than that of normal hemoglobin.

4.2.8 Estimation of Specific Heat Capacity

Specific heat capacity of hemoglobin protein is another important parameter to understand the transportation of heat through the human body. The thermal conduction of protein is the combine effects of density, specific heat capacity and thermal diffusion. The thermal diffusivity of protein measure, how quickly the temperature will change when the protein is heated or cooled. Proteins with a higher thermal diffusivity will heated or cooled quickly and vice versa. We have taken the systems of sickle and normal hemoglobin for estimating of their thermal parameters like the specific heat capacity by using the Equation 3.4.

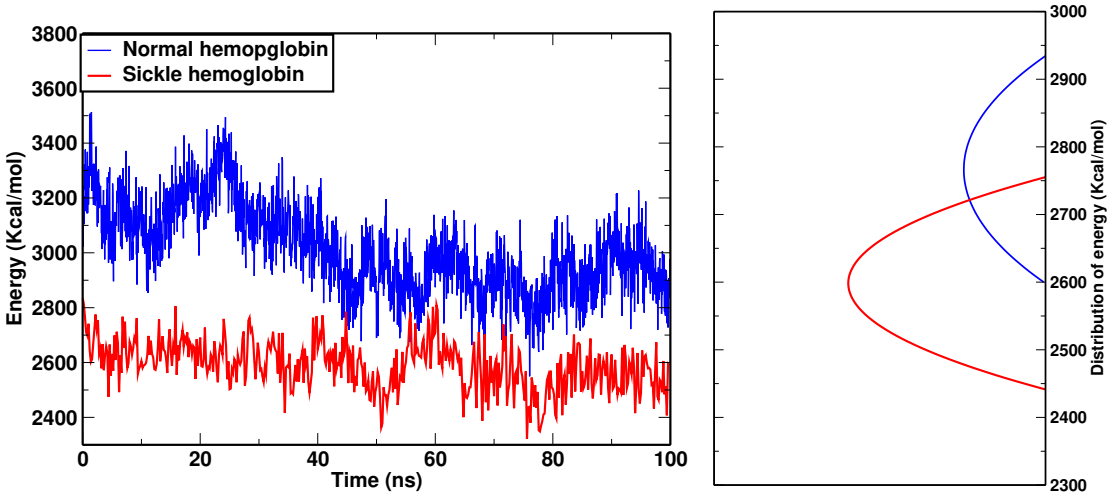


Figure 57: (a) Total energy fluctuation of sickle and normal hemoglobin protein at 310 K and (b) distribution of energy (Powrel & Adhikari, 2021).

Figure 57 (a) and (b) are the representations of the fluctuation and distribution of energy during NVT run. They explain the deviation of energy from the mean energy and distribution pattern of energy. More energy represents more velocity as well as the temperature of the system of particles. The square of the average total energy and average of the square of the total energies of the different systems are used in estimating the specific heat capacities (Table 13). On putting these values in to Equation 3.4, the specific heat capacities of sickle hemoglobin protein and normal hemoglobin protein are

found as 1966 Jkg^{-1-1} and 3204 Jkg^{-1-1} (Table 13).

Table 13: For normal and sickle cell hemoglobin protein at 310K of temperature(Powrel & Adhikari, 2021).

For	E_{Av}	$\langle E \rangle^2$	$\langle E^2 \rangle$	C_v $(J\ kg^{-1}K^{-1})$
Hb of	kcal mol^{-1}	(kcal $mol^{-1})^2$	(kcal $mol^{-1})^2$	$(\langle E^2 \rangle$-$\langle E \rangle^2)/k_B T^2$
Normal	2590.90	6712781.95	6721493.40	3204
Sickle	2792.83	7799890.27	7805237.47	1966

For the estimation of specific heat capacity at constant volume, we have used some of the theoretical and suitable simulation parameters as per classical molecular dynamics: k_B =0.00198657 kcal mol^{-1} K^{-1} , $\sigma \approx 8$ K, T=310 K and J=1.43846×10^{20} kcal/mol (Powrel & Adhikari, 2021). The specific heat capacity of sickle hemoglobin (1966 J/(kg-K)) (Table 13) is less than that of normal hemoglobin (3204 J/(kg-K)).

4.2.9 Diffusion of Sickle and Normal Hemoglobin Dimer

Hemoglobin diffusion has been considered as the means for oxygen diffusion through a solution (Longeville & Stingaciu, 2017). We have studied the diffusion of sickle and normal hemoglobin protein. In this subsection we have presented the self and binary diffusion of sickle and normal hemoglobin protein dimer. We have also estimated the temperature dependency of diffusion within the range of 306 K to 312 K using GROMACS (Berendsen et al., 1995; Lindahl et al., 2010; Lee et al., 2016). This work is focused in the estimation of diffusion coefficients near the human body temperature. The MSD graph are plotted for estimating the self diffusion coefficients of sickle, normal hemoglobin and water. During diffusion, the atoms are contagiously moving away from each other. The path followed by atoms during diffusion is expressed with radial distribution functions. It explains the structural and diffusive properties of the protein molecule. Einstein's equation 3.8 is implemented to find the self-diffusion coefficient. Simulation were carried out for the different temperatures up to 50 ns. Figures 58 represents the MSD versus time plot for HbS at 306 K, 308 K, 310 K and 312 K temperatures. It shows that as the temperature increases, the slope of MSD graph is also increased, which in turn raises the self-diffusion coefficient. We have shown the estimated values of self-diffusion coefficients of sickle, normal hemoglobin and water. These values are compared with the previously reported experimental results.

MSD of Sickle and Normal Hemoglobin Proteinat at Different Temperatures

For diffusion of protein in aqueous solution, we have used the GROMACS software for estimation of the MSD of atoms in the system of both sickle and normal hemoglobin dimer at temperatures of 306 K, 308 k, 310 K and 312 K. The slope of the MSD graph

are tabulated and calculated the self diffusion coefficients at required temperature. In order to understand mass transfer rate, we have studied diffusion of both sickle and normal hemoglobin dimer in TIP3P water model at given temperature. The diffusion is measured in terms of diffusion coefficient which can be measured using Einstein's relation (Equation 3.8). For this, we have plotted the MSD versus time graph for both solute and solvent molecules. The graphs are linearly fitted as shown in figures (Figure 58), Figure 59 and Figure 60. From the slope of the linearly fitted graphs, the estimated values of self diffusion coefficients as well as their binary diffusion coefficients are presented in Tables 14 and 15. We have also compared our estimated values with previously reported values.

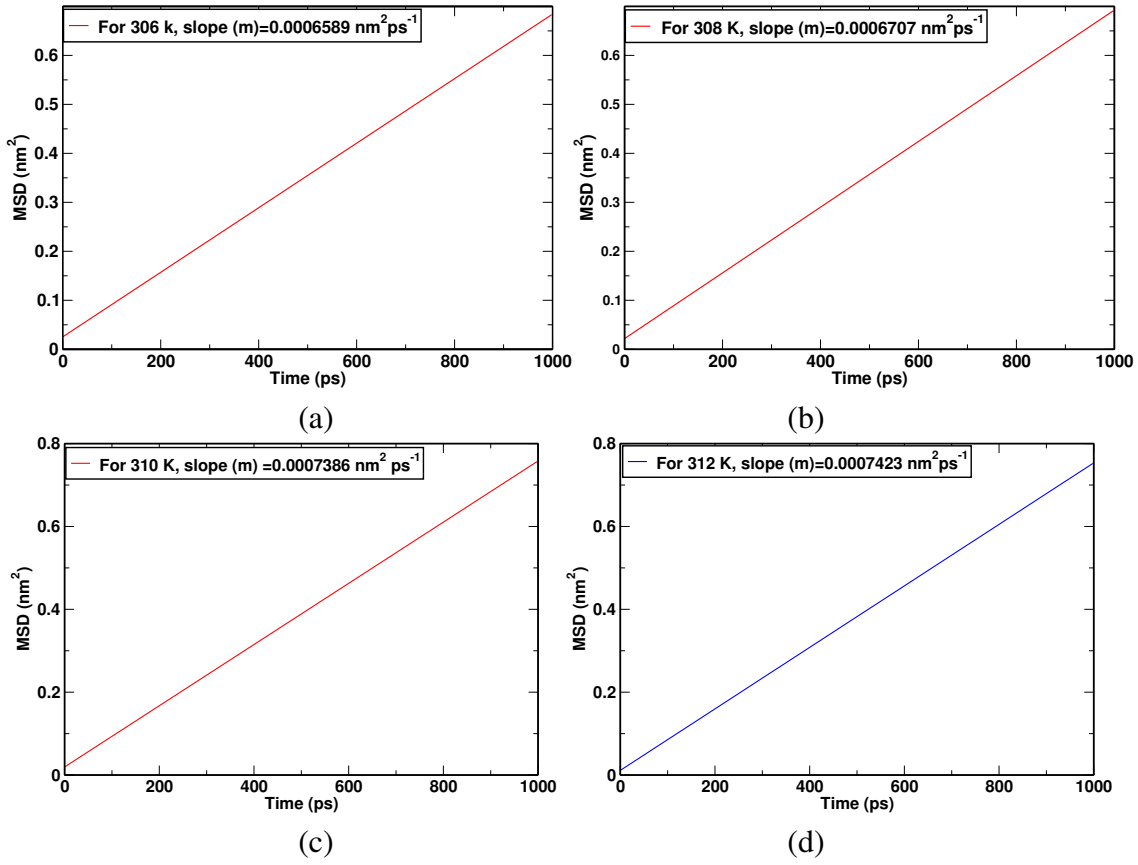


Figure 58: MSD versus time plots of sickle hemoglobin at (a) 306 K, (b) 308 K, (c) 310 K and (d) 312 K.

The Figure 58 represents the MSD graph of sickle hemoglobin dimer for obtaining the self diffusion coefficients. The Table 14 shows the slope of MSD graph of sickle hemoglobin at temperatures of 306 K, 308 K, 310 K and 312 K. The Table 16 contains the theoretically calculated and some of the available experimental values of the diffusion coefficients. The computational estimated values of sickle Hb are in the range of experimental finding values of diffusion.

Table 14: MSD of sickle hemoglobin at temperatures of 306 K, 308 K, 310 K and 312 K.

For	Slope ($nm^2 ps^{-1}$)		Diffusion ($\times 10^{-10}$ $m^2 s^{-1}$)	
Temp.(K)	Sickle Hb	Water	Sickle Hb	Water
306	0.0006580	0.031081	1.09	51.80
308	0.0006707	0.032564	1.11	54.27
310	0.0007386	0.033344	1.23	55.56
312	0.0007423	0.034427	1.24	57.37

Figures 59 shows the MSD vs time plot for HbA at 306 K, 308 K, 310 K 312 K temperatures. The slope of the graph gives the diffusion coefficient (D) of the normal hemoglobin protein at given temperatures.

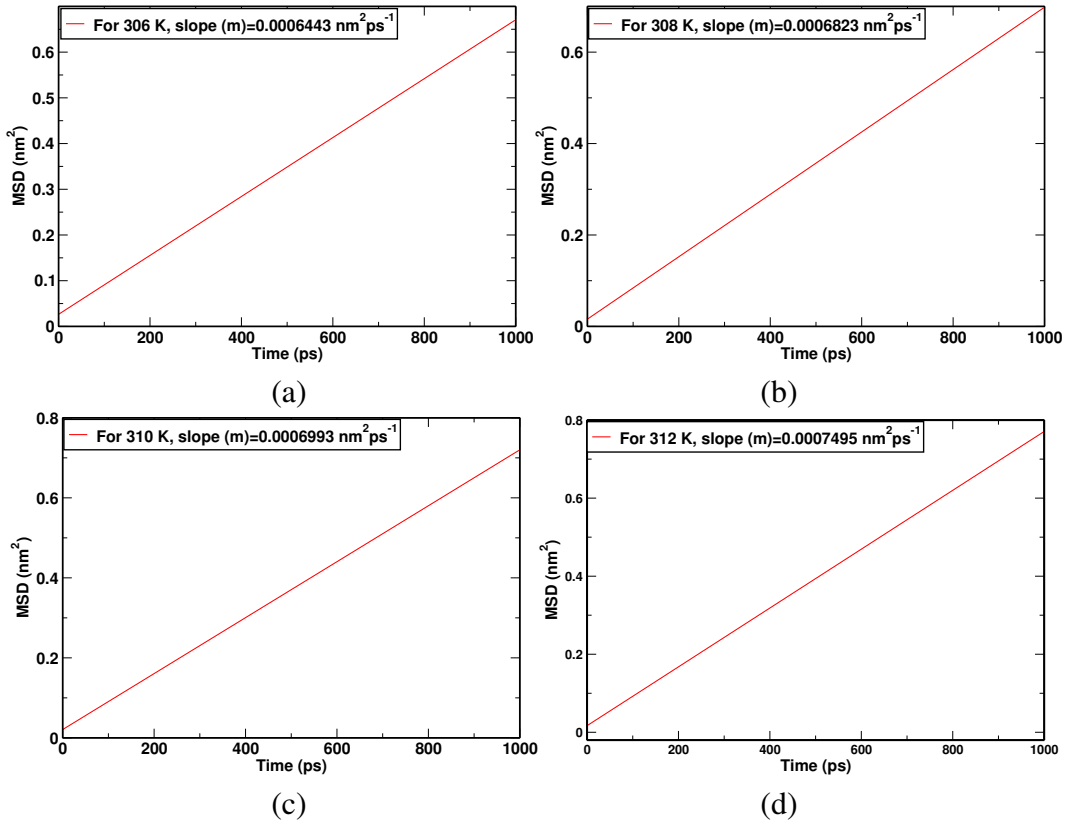


Figure 59: MSD versus time plots of normal hemoglobin at (a) 306 K, (b) 308 K, (c) 310 K and (d) 312 K.

The slope of the MSD graph (Figure 59) is used for estimating the diffusion constants of the normal hemoglobin protein as in Table 15. Here are the plotss of MSD versus time for water at 306 K, 308 K, 310 K, and 312 K temperatures. These graphs provides the diffusion coefficients of water (TIP3P model).

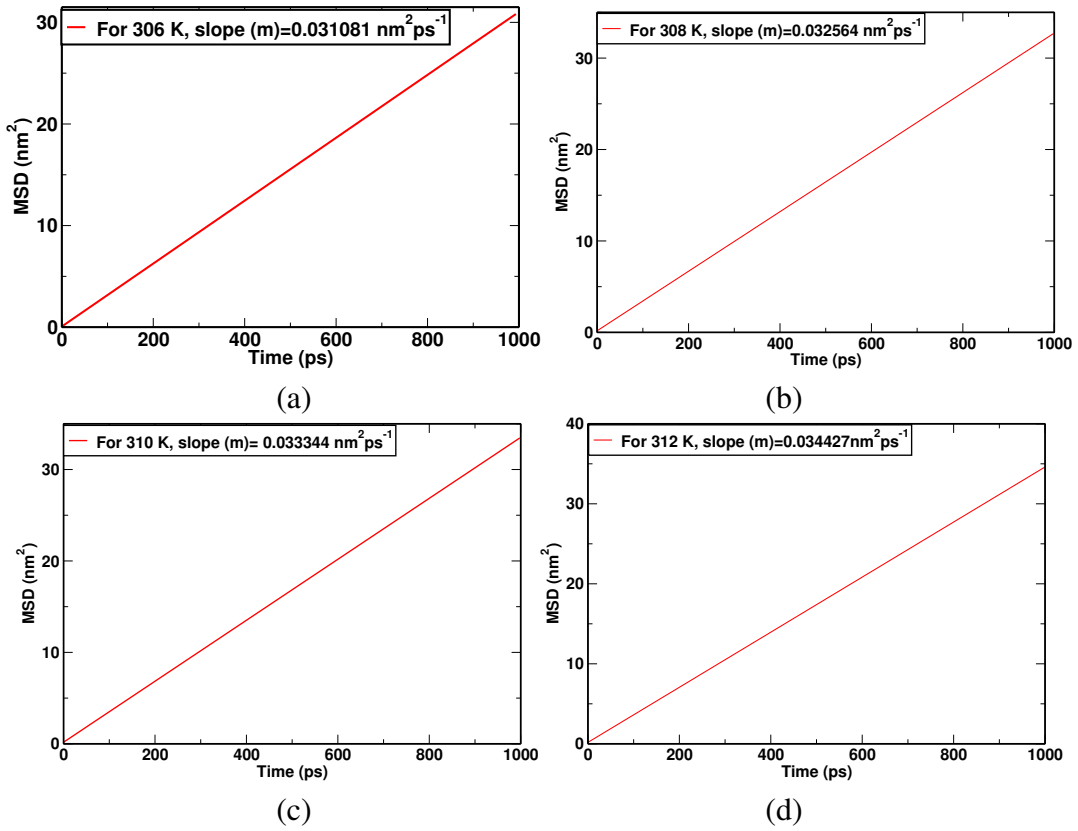


Figure 60: MSD versus time plots of water with hemoglobin protein at (a) 306 K, (b) 308 K, (c) 310 K and (d) 312 K..

The slope of the MSD graph of water are tabulated so as to estimates the diffusion coefficients of water as in Table 14 (sickle) and Table 15 (normal).

Table 15: Slope of MSD graph of normal hemoglobin at temperature of 306 K, 308 K, 310 K and 312 K.

For	Slope (nm^2 ps^{-1})		Diffusion ($\times$ 10^{-10} m^2 s^{-1})	
Temp.(K)	Normal Hb	Water	Normal Hb	Water
306	0.0006443	0.031081	1.07	51.80
308	0.0006823	0.032564	1.14	54.27
310	0.0006993	0.033344	1.16	55.56
312	0.0007495	0.034427	1.25	57.37

The sets of graphs in Figure (58), Figure (59) and Figure (60) are the plots within 10 ns of simulation time. This range covers the initial kinds of motion of the particles, showing a regular pattern. The initial straight portion of the graph is considered to be free from noise. Above 10 ns of simulation time, each of the graphs showed a different nature due to damping of the motions. The table includes the slopes of the graphs of r^2 versus time (t).

Effect of Temperature on Diffusion Coefficient

The diffusion of sickle and normal hemoglobin in TIP3P water varies with temperature. The results are tabulated in Table 16. It is seen that the variation of diffusion coefficient occurs with

the variation in temperature.

Table 16: Temperature dependency of diffusion Coefficient of TIP3P water, sickle and normal hemoglobin Proteins.

Temperature (K)	Diffusion coefficients (10^{-10} m^2 s^{-1})				
	D_{self} of HbS	D_{self} of A3N	D_{self} of Water	Binary Diff.(D_B) of HbS	Binary Diff.(D_B) of A3N
306	1.09	1.07	51.80	51.78	51.79
308	1.11	1.14	54.27	54.25	54.35
310	1.23	1.16	55.56	55.57	55.58
312	1.24	1.25	57.37	57.35	57.54

The experimentally determined diffusion coefficient of blood at 310 K is 7 $\times 10^{-11}$ m$^2 s^{-1}$ (Rugh & Bharathan, 2005). It is observed that as the temperature of both proteins increased from 306 K to 310 K, their diffusion rates also increased. However, above body temperature of 310 K the diffusion rate does not change smoothly due to the high temperature noises in the systems (Table (16)). It is seen from the Table (16) that the theoretical estimated self-diffusion coefficients are in the range obtained from the experimental results with minimum percentage error. It is seen from the Table 16 that the binary diffusion coefficient of normal is greater than that of sickle hemoglobin protein.

4.2.10 Binding Free Energy by Umbrella Sampling

A hemoglobin molecule has four peptide chains of molecules bonded with each other by specific amount of energy. Such a hemoglobin structure has two $\alpha\beta$ dimers. The single $\alpha\beta$ dimer of sickle and normal hemoglobin are shown in the Figures 61 (a) and (b) respectively. Sum of the interaction energy of the different residues in the protein chain creates a net binding energy affinity in the chain to bound with the other chains. We have focused on the estimation of the binding free energy in the $\alpha\beta$ dimer of sickle and normal hemoglobin proteins.

For this study, we have taken the structure of oxygenated (R-state) carbonmonoxyl-ligated sickle hemoglobin (HBS) which does not form the polymer in HBS (Ghatge et al., 2016b; M. Perutz et al., 1968; M. F. Perutz & Lehmann, 1968). Then, we investigated the free energy profile and atomic level interaction between the $\alpha\beta$ chains of the dimer. We have also studied the normal hemoglobin dimer structure by making the change in the mutated part of the already solved sickle hemoglobin dimer using CHARMM-GUI. We have carried out umbrella sampling simulation run for the study of binding free energy. Umbrella sampling molecular dynamics (MD) simulation is the process implemented to understand the structural conformation and stability of the normal and sickle hemoglobin dimer. The results of umbrella sampling of sickle protein and normal hemoglobin have been compared to the R-state conformation of carbonmonoxy. For this purpose of umbrella sampling analysis, WHAM is used in embedded NAMD environment. The binding free energy of sickle and normal hemoglobin is shown in Figure 62.

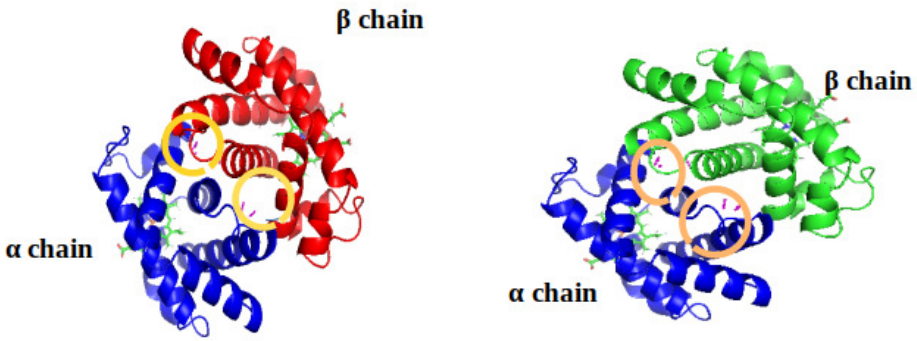


Figure 61: Binding in (a) sickle HbS and (b) normal hemoglobin dimer (Powrel et al., 2022).

The mutated beta chain in the sickle hemoglobin dimer interacts with the alpha chain by only three polar bonds as shown in Figure (61) (a). However, there are seven polar bonds in normal hemoglobin which bind the beta chain with the alpha chain (Figure (61) (b)).

Using the umbrella examining method, the free energy between the chains in sickle and normal hemoglobin has been calculated. For this, setup were arranged (as mentioned in system setup of dimer in subsection 3.4.1) for each window by taking the molecular separation of 1 Å along center of mass (COM) (Koirala, Pokhrel, et al., 2021). Beta chain (PROB) of each complex were pulled in positive x direction taking the alpha chain (PROA) as the reference chain. Twelve samples were prepared for sickle and normal hemoglobin dimer separately. Each of the samples are then executed for 20 ns simulation run under NPT condition.

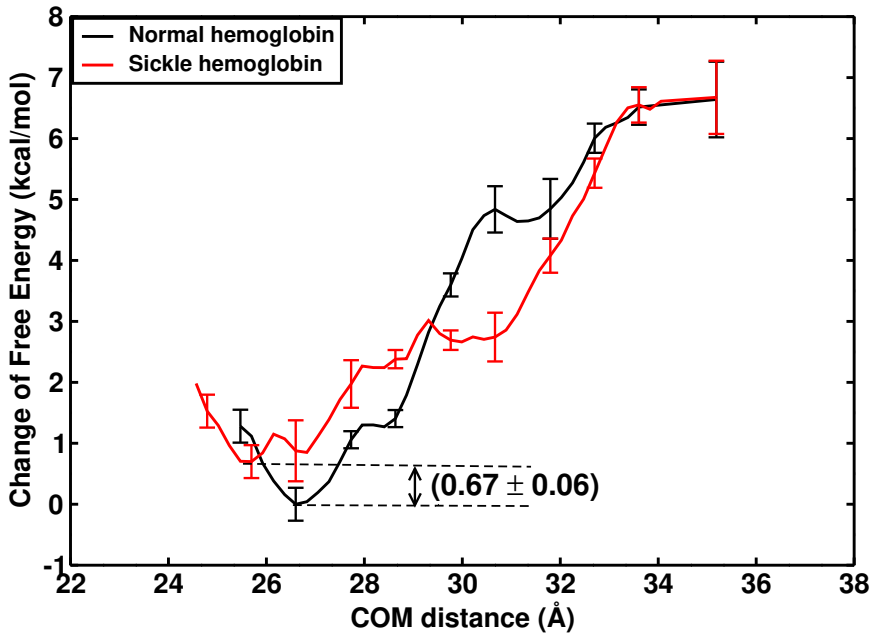


Figure 62: Free energy graph of sickle and normal hemoglobin proteins (Powrel et al., 2022)

.

The Figure 62 is the representation of the free energy difference of the beta chain from alpha in both sickle (red) and normal (black) hemoglobin. The graphs are so adjusted that at the

end point window the free energy of both the structure remains stable. It is done by re-scaling the vertical axis and compared the free energy difference at the pitfall point of the graphs. To find the validity of the data, error bars are plotted. Non overlapped small error bars at the different COM distance indicates the validity of the data. There is not any overlapping of the error bars in the two plots (Figure 62) at the interested comparing zone, which ensures that the free energies are significantly different in sickle and normal hemoglobin dimer. The graph shows higher difference of free energy for normal hemoglobin than that of sickle. It implies the favorable bindings of normal hemoglobin. Estimation of free energy profile have shown an energy difference of (5.97±0.27) kcal/mol in sickle dimer and (6.64±0.27) kcal/mol in normal hemoglobin dimer. These two free energies are now used to obtain the difference between them and their standard error.

We have estimated the standard deviation of the free energy at the pitfall point of the free energy graph of sickle ($\triangle G_S$) and normal hemoglobin ($\triangle G_N$) protein. The corresponding values of the free energy at the pitfall points are G_S and G_N . From our study, the binding free energy difference of normal and sickle hemoglobin dimer is found to be (0.67±0.06) kcal/mol (Powrel et al., 2022). This observation has shown the binding energy of $\alpha - \beta$ dimer in sickle is less than that of normal hemoglobin dimer by (0.67±0.06) kcal/mol (Powrel et al., 2022). It is due to the higher vdW and electrostatic interactions between chain α and chain β of the dimer. This result is also consistent with the higher number of residue pairs involved in the formation of the salt bridge and H-bonds in normal hemoglobin than in sickle. The study of the hydrophobic interaction has shown a higher number of residue pairs involved in the hydrophobic interaction of sickle dimer protein than in normal hemoglobin. They also cause a greater contact area between chains A and B of the sickle than normal hemoglobin dimer. This study realized that the alpha-beta dimer in normal hemoglobin has a stronger binding affinity than that of the sickle hemoglobin protein.

CHAPTER 5

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 General Consideration

The hemoglobin protein and dimer structures of sickle and normal hemoglobin proteins are studied using molecular dynamics. In hemoglobin structure, we have studied SMD of alpha and beta chains for finding the forces required in breaking the hydrogen bonds and Coloumbic interactions present in them. To make a depth understanding of the binding mechanism we further studied binding free energy between alpha and beta chains in the dimer of sickle and normal hemoglobin. The physical parameters of specific heat capacity and diffusion coefficients are also estimated and then compared with the experimental values. The comparison of the data showed good agreement with the available experimental data.

5.1.1 Elastic Property of Sickle and Normal Hemoglobin Tetramer Protein

The SMD simulation of the beta chain has shown that the structure of the sickle and de-oxygenated normal hemoglobin achieves its stable form after 20 ns and 25 ns from the start of the simulation. The RMSD values were also observed for them. Sickle protein becomes stable at 0.036 nm and normal hemoglobin at 0.035 nm. These values indicate the less flexibility of the atoms in the backbone of both sickle and normal hemoglobin protein chains. The β chain is found strongly interacted with hydrogen bonds and hydrophobic interaction. Hydrogen bonds are found higher in number in sickle than in normal hemoglobin. The vdW and electrostatic interaction energy of sickle protein have been observed more than that of the other three chains. Our investigation on hemoglobin protein finds higher hydrophobic interaction in sickle than that in normal hemoglobin. The salt bridges are also estimated stronger in normal hemoglobin than in sickle hemoglobin. Eight residue pairs are mainly contributing to the salt bridges in both proteins. Among the residues involved, lysine and asperin residues have a stronger contribution to the salt bridge formation in normal hemoglobin. At the same time, the contribution of these residues is reduced in sickle hemoglobin protein. On the contrary, glutamic acid and lysine residues contribute to the salt bridges of sickle hemoglobin proteins.

The SMD is also suitable for confirming the forces required for breaking the H-bonds in the protein chain. They indicate the stiffness of the proteins which depends on the values of the spring constant k used in SMD. Our investigation has shown higher stiffness in both sickle and

normal hemoglobin tetramer for higher values of spring constant and pulling velocities. The hydrophobic nature and the estimated SASA of protein in their solution are found somehow correlated. As the SASA of normal hemoglobin is much less than that of sickle hemoglobin protein so, we found a less hydrophilic nature in the normal protein than in sickle hemoglobin tetramer and dimer. Sticking together of the protein molecules is due to the hydrophilic nature of the residues. It causes in the increase in SASA of the hemoglobin protein molecules. The forces required in breaking the hydrogen bonds of the alpha chain with the other three chains in the sickle hemoglobin protein are found to be 6780 pN to 12345 pN with pulling velocities of 0.00020 nm/ps to 0.00040 nm/ps. A much higher force of 8738 pN to 16557 pN is required for breaking the hydrogen bonds of the alpha chain in normal hemoglobin. i.e. the beta chain have shown greater stiffness in comparison to alpha chain of both sickle and normal Hb proteins. Our study reveals stronger H-bonding, salt bridge, vdW and electrostatic interactions hydrophobic interaction, SASA in sickle hemoglobin protein than in normal hemoglobin protein. This result indicates that the alpha and beta chain of normal hemoglobin has higher stiffness than the alpha and beta chain of sickle hemoglobin at a temperature of 310 K. Among them beta is stiffer. It is in supports the theoretical results of higher free energy in normal hemoglobin protein than in sickle one.

It is thus concluded that the hydrophobic interactions are the intermediate factors that tend to guide the increase in the number of hydrogen bonding in the protein of sickle and normal Hb. Hence active hydrophobic interactions of the sickle can cause the sickling of the hemoglobin.

5.1.2 Free Energy of Sickle and Normal Hemoglobin Dimer Proteins

The binding energy of alpha-beta dimer of R-state sickle and normal hemoglobin protein are estimated by umbrella sampling molecular dynamics. The sickle dimer having valine at its sixth position estimates the free energy of 5.97 kcal/mol. The normal hemoglobin with glutamic acid at sixth position have shown the free energy of 6.64 kcal/mol. The free energy difference of normal and sickle hemoglobin protein dimer are compared for knowing the binding mechanism in them. The binding mechanism has also been analyzed by investigating the hydrogen bonding, hydrophobic interaction, salt bridges, electrostatics and van-der Waal's interactions. The alpha-beta dimer in normal has higher binding free energy than that of sickle hemoglobin dimer. After observing binding energy we have estimated the solvent accessible surface area (SASA) of the both dimer system. Our result have shown greater SASA in sickle dimer than that of normal hemoglobin dimer. We have also got the greater SASA in sickle tetramer structure than of normal hemoglobin dimer. Greater SASA results the much higher contact area at the interface of two chains (alpha-beta) in the sickle dimer than normal. The higher hydrophobic interaction in sickle dimer suppresses the sickle protein to have significantly higher contact area of two chains in the dimer and hence the sickling may appear. The free energy difference by umbrella sampling is observed as $\sim$ (6.64$\pm$0.27) kcal/mol in normal hemoglobin. For sickle hemoglobin it is found as $\sim$ (5.97$\pm$0.27) kcal/mol. The free energy of normal is higher by (0.67$\pm$ 0.06) kcal/mol. The lower free energy difference in sickle hemoglobin may be the cause of sickle shape hemoglobin

protein.

5.1.3 Specific Heat Capacity of Sickle and Normal Hemoglobin Proteins

We have investigated the specific heat capacity of sickle and normal hemoglobin protein. Our estimated specific heat capacity of sickle hemoglobin protein is 1966 J/kg-C at the temperature of 310 K. However for normal hemoglobin it is 3204 J/kg-C, which is greater than that of sickle hemoglobin. The inter-relationship of the estimated thermal and transport parameters with elastic properties are further used to make the depth knowledge of anti sickling drugs.

Studies on the thermodynamic characteristics of sickle and normal hemoglobin proteins have been conducted within the context of conventional molecular dynamics. According to our research, normal hemoglobin protein has a higher specific heat capacity than that of sickle hemoglobin protein.

5.1.4 Diffusion of Sickle and Normal Hemoglobin

MD simulations were used to estimate the diffusion rates of sickle and normal hemoglobin proteins at various temperatures. The results showed that the self diffusion coefficient and binary diffusion coefficients matched experimental findings. The diffusion rate of blood in normal hemoglobin was found to be slower than in sickle hemoglobin. Diffusion was measured in terms of self diffusion coefficients of solute and solvent, as well as their binary diffusion coefficient, at temperatures of 306 K, 308 K, 310 K, and 312 K. The diffusion rate of both hemoglobin proteins was found to be higher at 310 K. The results indicated that the diffusion coefficient of sickle protein was lower than that of normal hemoglobin protein dimer. Both self and binary diffusion coefficients increased with temperature, and the effect of temperature on self diffusion coefficients was analyzed using the Arrhenius equation. The estimated values of self diffusion coefficients followed Arrhenius behavior.

Recommendations

The binding mechanism of proteins or peptide chains in the hemoglobin protein complex is an important biophysical phenomenon. The non-equilibrium NC terminal SMD simulations can explain the dissociation mechanism and binding process of peptide chains. The SMD of one of the chains explains the unfolding way from the remaining three peptide chains. We thus recommend the techniques of SMD as suitable for the investigation of the elastic property of the protein. The interaction potential of protein molecules affects the transport and inter facial tension between the dimer of the hemoglobin protein. Hence, inter-facial tension is essential to estimate. The binding energy of alpha-beta dimer in normal hemoglobin is greater than that of sickle hemoglobin. It indicates that the binding energy is lost during the sickling of the hemoglobin protein. Thus, to recover the sickling of protein, energy must be added to the sickle Hb protein. This can be done by genetic engineering treatment. In this process HBB gene producing mutated residue of sickle cell is replaced by the original one for preventing the genetic disease like sickle cell.

The study of specific heat capacity and mass diffusion of hemoglobin indicates that the drug material must have the property for raising the specific heat capacity of sickle protein to become normal. As a result, our findings imply that a potential antisickling medication might be able to reduce the formation of a few salt bridges in between the α and β chains of sickle protein. Further study of molecular dynamics with docking utilizes these estimated parameters to develop suitable drugs.

CHAPTER 6

6. SUMMARY

In this work, we have studied the structure, elastic property, thermodynamics and binding mechanism of sickle and normal hemoglobin protein using molecular dynamics. Steered molecular dynamics (SMD) and umbrella sampling (US) techniques are implemented to understand the elastic property and binding energy in the complexes. The elastic property of beta and alpha chains in both the sickle and normal hemoglobin tetramer proteins are estimated. Normal hemoglobin was found more elastic than sickle hemoglobin protein. Our study on non-bonded interaction shows that the interactions of the beta chain in both hemoglobin have a higher contribution than that of the alpha chain, which makes the sickle shape to the Hb protein. We estimated the free energy of normal hemoglobin higher than that of sickle hemoglobin. Also, we have estimated the specific heat capacity of both the hemoglobin proteins. Our estimated values are in close agreement with the previously reported experimental values. To understand the transport phenomenon and the effect of temperature on the diffusion of the hemoglobin proteins, we have performed the simulations at four different temperatures: 306 K, 308 K, 3010 K and 312 K. It has observed that the diffusion coefficient for both solute and solvent follow Arrhenious behavior. We have also found the specific heat capacity of normal hemoglobin higher than that of sickle hemoglobin. The coefficient of mass diffusion of sickle is lower than that of normal hemoglobin protein. Moreover, the time dependency of mass diffusion has been studied for both the sickle and normal hemoglobin dimer structures.

The results of umbrella sampling also support the stronger interaction between the alpha-beta chain of normal hemoglobin than that of the sickle. The normal hemoglobin dimer structure exceeds the binding energy of sickle hemoglobin dimer by (0.67±0.06) kcal mol^{-1}.

It thus recommends to increase the binding force of dimer and diffusive property of sickle protein to recover into normal form. Although, the sickle is stronger because of higher vdW and electrostatics interactions energy than that of normal Hb, which may be due to higher ion concentration in the molecules. But hydrogenation of the sickle protein may reduce ion concentration as well as increase salt bridges and the number of hydrogen bonds to recover the sickling.

The recommended drug should have the ability to increase the binding affinity of dimer and diffusive property of Hb Protein.

REFERENCES

Abdollahzadeh Jamalabadi, M., Akbari Bidokhti, A. A., Khak Rah, H., Vaezi, S., & Hooshmand, P. (2016). Numerical investigation of oxygenated and deoxygenated blood flow through a tapered stenosed arteries in magnetic field. *PloS one*, *11*(12), e0167393.

Abdollahzadeh Jamalabadi, M., Ebrahimi, M., Homayoun, G., & Hooshmand, P. (2015). Influence of radiative heat transfer and transverse magnetic field on peristaltic flow of a third order fluid in a planar channel. *Journal of Chemical and Pharmaceutical Research*, *7*(12), 788–799.

Abdollahzadeh Jamalabadi, M., & Keikha, A. (2016). Fluid-solid interaction modeling of cerebrospinal fluid absorption in arachnoid villi. *Journal of Chemical and Pharmaceutical Research*, *8*(2), 428–442.

Akbar, N., Khan, Z., & Nadeem, S. (2016). Influence of magnetic field and slip on jeffrey fluid in a ciliated symmetric channel with metachronal wave pattern. *Journal of Applied Fluid Mechanics*, *9*(2), 565–572.

Akbar, N. S. (2016). Metallic nanoparticles analysis for the blood flow in tapered stenosed arteries: Application in nanomedicines. *International Journal of Biomathematics*, *9*(01), 1650002.

Akbar, N. S., & Butt, A. W. (2016). Entropy generation analysis for metachronal beating of ciliated cu-water nanofluid with magnetic field. *International Journal of Exergy*, *19*(1), 41–54.

Akbar, N. S., Raza, M., & Ellahi, R. (2016). Copper oxide nanoparticles analysis with water as base fluid for peristaltic flow in permeable tube with heat transfer. *Computer Methods and Programs in Biomedicine*, *130*, 22-30. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0169260715303035` doi: https://doi.org/10.1016/j.cmpb.2016.03.003

Akinsheye, I., & Klings, E. S. (2010). Sickle cell anemia and vascular dysfunction: the nitric oxide connection. *Journal of cellular physiology*, *224*(3), 620–625.

Alder, B., & Wainwright, T. (1967). Velocity autocorrelations for hard spheres. *Physical Review Letters*, *18*(23), 988.

Alder, B. J., & Wainwright, T. E. (1957). Phase transition for a hard sphere system. *The Journal of Chemical Physics*, *27*(5), 1208–1209.

Alexy, T., Pais, E., Armstrong, J. K., Meiselman, H. J., Johnson, C. S., & Fisher, T. C. (2006). Rheologic behavior of sickle and normal red blood cell mixtures in sickle plasma: implications for transfusion therapy. *Transfusion*, *46*(6), 912–918.

Allen, M., & Tildesley, D. (1987). *J. computer simulations of liquids 1987*. Oxford University Press, New York.

Allen, M. P., & Tildesley, D. J. (2017a). *Computer simulation of liquids*. Oxford university press.

Allen, M. P., & Tildesley, D. J. (2017b). *Computer Simulation of Liquids*. Oxford University Press. Retrieved from `https://doi.org/10.1093/oso/9780198803195.001.0001` doi: 10.1093/oso/9780198803195.001.0001

Allison, A., et al. (1956). Observations on the sickling phenomenon and on the distribution of different haemoglobin types in erythrocyte populations. *Clinical Science*, *15*(4), 497–510.

Andersson, M., Bennetzen, M., Klamt, A., & Stipp, S. S. (2014). First-principles prediction of liquid/liquid interfacial tension. *Journal of chemical theory and computation*, *10*(8), 3401–3408.

Antonini, E., & Brunori, M. (1970). Hemoglobin. *Annual review of biochemistry*, *39*(1), 977–1042.

Arnone, A. (1974). Mechanism of action of hemoglobin. *Annual review of medicine*, *25*(1), 123–130.

Bailey, K. (1948). Fibrous proteins as components of biological systems. *British Medical Bulletin*, *5*(4-5), 338–341.

Banach, M., Konieczny, L., & Roterman, I. (2019). Secondary and supersecondary structure of proteins in light of the structure of hydrophobic cores. In *Protein supersecondary structures* (pp. 347–378). Springer.

Barpanda, S. S. (2013). *Use of image processing techniques to automatically diagnose sickle-cell anemia present in red blood cells smear* (Unpublished doctoral dissertation).

Barrick, D., Ho, N. T., Simplaceanu, V., Dahlquist, F. W., & Ho, C. (1997). A test of the role of the proximal histidines in the perutz model for cooperativity in haemoglobin. *Nature structural biology*, *4*(1), 78–83.

Berendsen, H. J., van der Spoel, D., & van Drunen, R. (1995). GROMACS: a message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, *91*(1-3), 43–56.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... Bourne, P. E. (2000). The protein data bank. *Nucleic acids research*, *28*(1), 235–242.

Billett, H. H. (1990). Hemoglobin and hematocrit. *Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition*.

Borgstahl, G. E. O. (1992). *High-resolution x-ray diffraction studies of hemoglobin: Ligand-induced tertiary structure movements in" locked" quaternary structures*. The University of Iowa.

Briehl, R. (1982). The effects of shear on the delay time for gelation of hemoglobin s. *Blood Cells*, *8*(2), 201–212.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. a., & Karplus, M. (1983). Charmm: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry*, *4*(2), 187–217.

Bussi, G., Zykova-Timan, T., & Parrinello, M. (2009). Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *The Journal of chemical physics*, *130*(7), 074101.

Callister, W. D. (2007). Materials science and engineering: An introduction john wiley & sons. *Inc.: Utah*.

Carr, R., & Cockelet, G. R. (1981). Rheology of suspensions of normal and hardened erythrocytes and their mixtures. *Journal of Rheology*, *25*(1), 67–82.

Carter, R., & Mendis, K. N. (2002). Evolutionary and historical aspects of the burden of malaria. *Clinical microbiology reviews*, *15*(4), 564–594.

Chakravarty, S., & Mandal, P. K. (2000). Two-dimensional blood flow through tapered arteries under stenotic conditions. *International Journal of Non-Linear Mechanics*, *35*(5), 779–793.

Charache, S., Terrin, M. L., Moore, R. D., Dover, G. J., Barton, F. B., Eckert, S. V., ... of the Multicenter Study of Hydroxyurea in Sickle Cell Anemia, I. (1995).

Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. *New England Journal of Medicine*, *332*(20), 1317–1322.

Crank, J. (1979). *The mathematics of diffusion*. Oxford university press.

Darken, L., et al. (2010). Diffusion, mobility and their interrelation through free energy in binary metallic systems. *Metallurgical and materials transactions. B, Process metallurgy and materials processing science*, *41*(2), 277–294.

DiStasio, R. A., Gobre, V. V., & Tkatchenko, A. (2014). Many-body van der waals interactions in molecules and condensed matter. *Journal of Physics: Condensed Matter*, *26*(21), 213202.

Dormandy, E., James, J., Inusa, B., & Rees, D. (2018). How many people have sickle cell disease in the uk? *Journal of Public Health*, *40*(3), e291–e295.

Eaton, W. A. (2002). Linus pauling and sickle cell disease. *Biophysical chemistry*, *100*(1-3), 109–116.

Eaton, W. A., Henry, E. R., Hofrichter, J., Bettati, S., Viappiani, C., & Mozzarelli, A. (2007). Evolution of allosteric models for hemoglobin. *IUBMB life*, *59*(8-9), 586–599.

Eaton, W. A., Hofrichter, J., & Ross, P. D. (1976). Delay time of gelation: a possible determinant of clinical severity in sickle cell disease.

El Hage, K., Hedin, F., Gupta, P. K., Meuwly, M., & Karplus, M. (2018). Valid molecular dynamics simulations of human hemoglobin require a surprisingly large box size. *Elife*, *7*, e35560.

Ercolessi, F. (1997). A molecular dynamics primer. *Spring college in computational physics, ICTP, Trieste*, *19*.

Evans, E., Mohandas, N., Leung, A., et al. (1984). Static and dynamic rigidities of normal and sickle erythrocytes. major influence of cell hemoglobin concentration. *The Journal of clinical investigation*, *73*(2), 477–488.

Ferrone, F. A., Hofrichter, J., & Eaton, W. A. (1985). Kinetics of sickle hemoglobin polymerization: Ii. a double nucleation mechanism. *Journal of Molecular Biology*, *183*(4), 611-631. Retrieved from `https://www.sciencedirect.com/science/article/pii/0022283685901755` doi: https://doi.org/10.1016/0022-2836(85)90175-5

Ferrone, F. A., & Rotter, M. A. (2004). Crowding and the polymerization of sickle hemoglobin. *Journal of Molecular Recognition*, *17*(5), 497–504.

Fischer, W.-N., Loo, D. D., Koch, W., Ludewig, U., Boorer, K. J., Tegeder, M., ... Frommer, W. B. (2002). Low and high affinity amino acid h+-cotransporters for cellular import of neutral and charged amino acids. *The Plant Journal*, *29*(6), 717–731.

Ghatge, M. S., Ahmed, M. H., Omar, A. S. M., Pagare, P. P., Rosef, S., Kellogg, G. E., ... Safo, M. K. (2016a). Crystal structure of carbonmonoxy sickle hemoglobin in r-state conformation. *Journal of Structural Biology*, *194*(3), 446-450. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1047847716300624` doi: https://doi.org/10.1016/j.jsb.2016.04.003

Ghatge, M. S., Ahmed, M. H., Omar, A. S. M., Pagare, P. P., Rosef, S., Kellogg, G. E., ... Safo, M. K. (2016b). Crystal structure of carbonmonoxy sickle hemoglobin in r-state conformation. *Journal of structural biology*, *194*(3), 446–450.

Giardina, B., Messana, I., Scatena, R., & Castagnola, M. (1995). The multiple functions of hemoglobin. *Critical reviews in biochemistry and molecular biology*, *30*(3), 165–196.

Gibson, J., Goland, A. N., Milgram, M., & Vineyard, G. (1960). Dynamics of radiation damage. *Physical Review*, *120*(4), 1229.

Gilson, M. K., Given, J. A., Bush, B. L., & McCammon, J. A. (1997). The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophysical journal*, *72*(3), 1047–1069.

González-Ruiz, D., & Gohlke, H. (2006). Targeting protein-protein interactions with small molecules: challenges and perspectives for omputational binding epitope detection and ligand finding. *Current medicinal chemistry*, *13*(22), 2607–2625.

Gulbis, B., Haberman, D., Dufour, D., Christophe, C., Vermylen, C., Kagambega, F., ... others (2005). Hydroxyurea for sickle cell disease in children and for prevention of cerebrovascular events: the belgian experience. *Blood*, *105*(7), 2685–2690.

Guo, W., Yuan, J., Dong, Q., & Wang, E. (2010). Highly sequence-dependent formation of fluorescent silver nanoclusters in hybridized dna duplexes for single nucleotide mutation identification. *Journal of the American Chemical Society*, *132*(3), 932–934.

Harris, J. W., & Bensusan, H. B. (1975). The kinetics of the sol-gel transformation of deoxyhemoglobin s by continuous monitoring of viscosity. *The Journal of Laboratory and Clinical Medicine*, *86*(4), 564–575.

Herrick, J. (1924). Abstract of discussion. *JAMA*, *83*, 16.

Herrick, J. B. (1910). Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. *Archives of internal medicine*, *6*(5), 517–521.

Ho, C., Willis, B. F., Shen, T.-J., Ho, N. T., Sun, D. P., Tam, M. F., ... Nagel, R. L. (1996). Roles of α114 and β87 amino acid residues in the polymerization of hemoglobin s: Implications for gene therapy. *Journal of molecular biology*, *263*(3), 475–485.

Hosseinzadeh, V. A., Brugnara, C., & Holt, R. G. (2018). Shape oscillations of single blood drops: applications to human blood and sickle cell disease. *Scientific reports*, *8*(1), 1–8.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38.

Ingram, V. M. (1956). A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature*, *178*(4537), 792–794.

Itoh, T., Chien, S., & Usami, S. (1995). Effects of hemoglobin concentration on deformability of individual sickle cells after deoxygenation.

Jia, Y., Duan, L., & Li, J. (2016). Hemoglobin-based nanoarchitectonic assemblies as oxygen carriers. *Advanced Materials*, *28*(6), 1312–1318.

Jin, W., & Chen, H. (2000). A new method of determination of diffusion coefficients using capillary zone electrophoresis (peak-height method). *Chromatographia*, *52*(1), 17–21.

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). Charmm-gui: a web-based graphical user interface for charmm. *Journal of computational chemistry*, *29*(11), 1859–1865.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935.

Kappel, K., Jarmoskaite, I., Vaidyanathan, P. P., Greenleaf, W. J., Herschlag, D., & Das, R. (2019). Blind tests of rna–protein binding affinity prediction. *Proceedings of the National Academy of Sciences*, *116*(17), 8336–8341.

Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature structural biology*, *9*(9), 646–652.

Kästner, J. (2011). Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(6), 932–942.

Kato, G. J., Steinberg, M. H., Gladwin, M. T., et al. (2017). Intravascular hemolysis and the pathophysiology of sickle cell disease. *The Journal of clinical investigation*, *127*(3), 750–760.

Kaul, D. K., Fabry, M., Windisch, P., Baez, S., Nagel, R., et al. (1983). Erythrocytes in sickle cell anemia are heterogeneous in their rheological and hemodynamic characteristics. *The Journal of clinical investigation*, *72*(1), 22–31.

Kaul, D. K., & Xue, H. (1991). Rate of deoxygenation and rheologic behavior of blood in sickle cell anemia. *Blood*, *77*(6), 1353–1361.

Koirala, R. P., Bhusal, H. P., Khanal, S. P., & Adhikari, N. P. (2020). Effect of temperature on transport properties of cysteine in water. *AIP Advances*, *10*(2), 025122.

Koirala, R. P., Pokhrel, R., Baral, P., Tiwari, P. B., Chapagain, P. P., & Adhikari, N. P. (2021). Structural insights into the repair mechanism of AGT for methyl-induced DNA damage. *Biological Chemistry*, *402*(10), 1203–1211.

Koirala, R. P., Thapa, B., Khanal, S. P., Powrel, J., Adhikari, R. P., & Adhikari, N. P. (2021). Binding of sars-cov-2/sars-cov spike protein with human ace2 receptor. *Journal of Physics Communications*, *5*(3), 035010.

Korotchkina, L. G., Ciszak, E. M., & Patel, M. S. (2004). Function of several critical amino acids in human pyruvate dehydrogenase revealed by its structure. *Archives of biochemistry and biophysics*, *429*(2), 171–179.

Krishna, R., & Van Baten, J. (2005). The darken relation for multicomponent diffusion in liquid mixtures of linear alkanes: An investigation using molecular dynamics (md) simulations. *Industrial & engineering chemistry research*, *44*(17), 6939–6947.

Krogh, A. (1941). The comparative physiology of respiratory mechanisms. In *The comparative physiology of respiratory mechanisms.* University of Pennsylvania press.

Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *Journal of computational chemistry*, *13*(8), 1011–1021.

Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., ... others (2016). Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Journal of chemical theory and computation*, *12*(1), 405–413.

Lennard-Jones, J. E. (1931). Cohesion. *Proceedings of the Physical Society (1926-1948)*, *43*(5), 461.

Levitt, M., & Warshel, A. (1975). Computer simulation of protein folding. *Nature*, *253*(5494), 694–698.

Li, H., & Lykotrafitis, G. (2011). A coarse-grain molecular dynamics model for sickle hemoglobin fibers. *Journal of the mechanical behavior of biomedical materials*, *4*(2), 162–173.

Lindahl, E., Hess, A., Buuren, R. v., Apol, E., Meulenhoff, P., Tieleman, D., ... Berendsen, H. (2010). *Gromacs user manual, version 4.5.6.*

Longeville, S., & Stingaciu, L.-R. (2017). Hemoglobin diffusion and the dynamics of oxygen capture by red blood cells. *Scientific reports*, *7*(1), 10448.

Luzzatto, L. (2012). Sickle cell anaemia and malaria. *Mediterranean journal of hematology and infectious diseases*, *4*(1).

Maciaszek, J. L., & Lykotrafitis, G. (2011). Sickle cell trait human erythrocytes are significantly stiffer than normal. *Journal of biomechanics*, *44*(4), 657–661.

Mann, F. C., Herrick, J., Essex, H. E., & Baldes, E. J. (1938). The effect on the blood flow of decreasing the lumen of a blood vessel. *Surgery*, *4*(2), 249–252.

Marcus, Y., & Hefter, G. (2006). Ion pairing. *Chemical reviews*, *106*(11), 4585–4621.

Marengo-Rowe, A. J. (2006). Structure-function relations of human hemoglobins. *Baylor University Medical Center Proceedings*, *19*(3), 239-245. Retrieved from `https://doi.org/10.1080/08998280.2006.11928171` (PMID: 17252042) doi: 10.1080/08998280.2006.11928171

McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature*, *267*(5612), 585–590.

Merkel, T. J., Jones, S. W., Herlihy, K. P., Kersey, F. R., Shields, A. R., Napier, M., ... others (2011). Using mechanobiological mimicry of red blood cells to extend circulation times of hydrogel microparticles. *Proceedings of the National Academy of Sciences*, *108*(2), 586–591.

Möller, D., Gellert, M., Langel, W., & Lillig, C. H. (2017). Molecular dynamics simulations and in vitro analysis of the crmp2 thiol switch. *Molecular BioSystems*, *13*(9), 1744–1753.

Monod, J., Wyman, J., & Changeux, J.-P. (1965). On the nature of allosteric transitions: a plausible model. *Journal of molecular biology*, *12*(1), 88–118.

Mosher, M. (1992). *Organic Chemistry. (Morrison, Robert Thornton; Boyd, Robert Neilson).* ACS Publications.

Muirhead, H., & Perutz, M. (1963). Structure of hæmoglobin: A three-dimensional fourier synthesis of reduced human haemoglobin at 5.5 å resolution. *Nature*, *199*(4894), 633–638.

Murcko, M. A. (1995). Computational methods to predict binding free energy in ligand-receptor complexes. *Journal of medicinal chemistry*, *38*(26), 4953–4967.

Murray, R. K., Granner, D. K., Mayes, P. A., & Rodweli, V. (1990). Harper's biochemistry 22nd ed. *Norwalk, CT: Appleton & Lange*, 499–510.

Noguchi, C. T., & Schechter, A. N. (1981). The intracellular polymerization of sickle hemoglobin and its relevance to sickle cell disease. *Blood*, *58*(6), 1057–1068.

Northrop, J. H., & Anson, M. (1929). A method for the determination of diffusion constants and the calculation of the radius and weight of the hemoglobin molecule. *The Journal of General Physiology*, *12*(4), 543.

Nosal-Wiercińska, A. (2012). Electrochemical and thermodynamic study of the electroreduction of bi (iii) ions in the presence of cysteine in solutions of different water activity. *Journal of Electroanalytical Chemistry*, *681*, 103–108.

Orkin, S. H., & Bauer, D. E. (2019). Emerging genetic therapy for sickle cell disease. *Annual review of medicine*, *70*, 257–271.

Pande, R., Ghimire, P. G., Chand, P. B., et al. (2019). Sickle cell disease in western nepal. *Nepal Journal of Medical Sciences*, *4*(1), 15–19.

Pauling, L. (1964). Molecular disease and evolution. *Bulletin of the New York Academy of Medicine*, *40*(5), 334.

Pauling, L., Itano, H. A., Singer, S. J., & Wells, I. C. (1949). Sickle cell anemia, a molecular disease. *Science*, *110*(2865), 543–548.

Pawliuk, R., Westerman, K. A., Fabry, M. E., Payen, E., Tighe, R., Bouhassira, E. E., ... others (2001). Correction of sickle cell disease in transgenic mouse models by gene therapy. *Science*, *294*(5550), 2368–2371.

Perutz, M., Muirhead, H., Cox, J., & Goaman, L. (1968). Three-dimensional fourier synthesis of horse oxyhaemoglobin at 2.8 å resolution: the atomic model. *Nature*, *219*(5150), 131–139.

Perutz, M. F. (1978). Hemoglobin structure and respiratory transport. *Scientific American*, *239*(6), 92–125.

Perutz, M. F., & Lehmann, H. (1968). Molecular pathology of human haemoglobin. *Nature*, *219*(5157), 902–909.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., . . . Schulten, K. (2005). Scalable molecular dynamics with namd. *Journal of computational chemistry*, *26*(16), 1781–1802.

Platt, O. S., et al. (2000). Sickle cell anemia as an inflammatory disease. *The Journal of clinical investigation*, *106*(3), 337–338.

Powrel, J. (2022). Computational estimation of free energy using fortran code. *Butwal Campus Journal*, *5*(1), 143–159. Retrieved from `https://nepjol.info/index.php/bcj/article/view/50199` doi: 10.3126/bcj.v5i1.50199

Powrel, J., & Adhikari, N. P. (2021). Thermal properties of normal and sickled hemoglobin protein. *BIBECHANA*, *18*(1), 140–148.

Powrel, J., & Adhikari, N. P. (2022). Elastic property of sickle and normal hemoglobin protein: Molecular dynamics. *AIP Advances*, *12*(4), 045308.

Powrel, J., Koirala, R. P., & Adhikari, N. P. (2022). Structural stability of r-state conformation of carbonmonoxyl sickle and normal hemoglobin dimer. *Journal of Biomolecular Structure and Dynamics*, *Advance online publication*, 1-10. Retrieved from `https://doi.org/10.1080/07391102.2022.2150890` (PMID: 36473709) doi: 10.1080/07391102.2022.2150890

Raha, K., & Merz Jr, K. M. (2005). Calculating binding free energy in protein–ligand interaction. *Annual reports in computational chemistry*, *1*, 113–130.

Rahman, A. (1964). Correlations in the motion of atoms in liquid argon. *Physical Review*, *136*(2A), A405.

Rees, D. C., Williams, T. N., & Gladwin, M. T. (2010). Sickle-cell disease. *The Lancet*, *376*(9757), 2018–2031.

Ribeil, J.-A., Hacein-Bey-Abina, S., Payen, E., Magnani, A., Semeraro, M., Magrin, E., . . . others (2017). Gene therapy in a patient with sickle cell disease. *New England Journal of Medicine*, *376*(9), 848–855.

Rugh, J. P., & Bharathan, D. (2005). Predicting human thermal comfort in automobiles. *SAE transactions*, 2508–2516.

Safo, M., & Abraham, D. (2005). The enigma of the liganded hemoglobin end state: a novel quaternary structure of human carbonmonoxy hemoglobin. *Biochemistry*, *44*(23), 8347–8359.

Schechter, A. N. (2008). Hemoglobin research and the origins of molecular medicine. *Blood, The Journal of the American Society of Hematology*, *112*(10), 3927–3938.

Schmalzer, E., Lee, J., Brown, A., Usami, S., & Chien, S. (1987). Viscosity of mixtures of sickle and normal red cells at varying hematocrit levels: implications for transfusion. *Transfusion*, *27*(3), 228–233.

Shrestha, A., & Karki, S. (2013). Analysis of sickle hemoglobin. *Journal of pathology of Nepal*, *3*(6), 437–440.

Stradner, A., Foffi, G., Dorsaz, N., Thurston, G., & Schurtenberger, P. (2007). New insight into cataract formation: Enhanced stability through mutual attraction. *Physical Review Letters*, *99*(19), 198103.

Thom, C. S., Dickson, C. F., Gell, D. A., & Weiss, M. J. (2013). Hemoglobin variants: biochemical properties and clinical correlates. *Cold Spring Harbor perspectives in medicine*, *3*(3), a011858.

Tina, K., Bhadra, R., & Srinivasan, N. (2007). Pic: protein interactions calculator. *Nucleic acids research*, *35*(suppl_2), W473–W476.

van Gunsteren, Wilfred F and Berendsen, Herman JC. (1988). A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, *1*(3), 173–185.

Van Eps, L. S., & De Jong, P. (1997). Sickle cell disease. *Diseases of the Kidney*, 2201–2219.

Virk, A. S., Stait-Gardner, T., Willis, S. A., Torres, A. M., & Price, W. S. (2015). Macromolecular crowding studies of amino acids using nmr diffusion measurements and molecular dynamics simulations. *Frontiers in Physics*, *3*, 1.

Weatherall, D. (2008). Genetic variation and susceptibility to infection: the red cell and malaria. *British journal of haematology*, *141*(3), 276–286.

Wells, I. C., & Itano, H. A. (1951). Ratio of sickle-cell anemia hemoglobin to normal hemoglobin in sicklemics. *Journal of Biological Chemistry*, *188*(1), 65–74.

Wishner, B. C., Ward, K., Lattman, E., & Love, W. (1975). Crystal structure of sickle-cell deoxyhemoglobin at 5 å resolution. *Journal of molecular biology*, *98*(1), 179–194.

Wong, C. F., & McCammon, J. A. (2003). Protein simulation and drug design. *Advances in protein chemistry*, *66*, 87–121.

Xiao, B.-L., Ning, Y.-N., Niu, N.-N., Li, D., Moosavi-Movahedi, A. A., Sheibani, N., & Hong, J. (2019). Steered molecular dynamic simulations of conformational lock of cu, zn-superoxide dismutase. *Scientific reports*, *9*(1), 1–10.

Xu, X., Rioux, T. P., & Castellani, M. P. (2022). The specific heat of the human body is lower than previously believed: The journal temperature toolbox. *Temperature*, *0*(0), 1-5. Retrieved from `https://doi.org/10.1080/23328940.2022.2088034` doi: 10.1080/23328940.2022.2088034

Xue, Y., Lofland, S., & Hu, X. (2019). Thermal conductivity of protein-based materials: A review. *Polymers*, *11*(3), 456.

Yesudasan, S., Douglas, S. A., Platt, M. O., Wang, X., & Averett, R. D. (2019). Molecular insights into the irreversible mechanical behavior of sickle hemoglobin. *Journal of Biomolecular Structure and Dynamics*, *37*(5), 1270–1281.

Yesudasan, S., Wang, X., & Averett, R. D. (2018). Molecular dynamics simulations indicate that deoxyhemoglobin, oxyhemoglobin, carboxyhemoglobin, and glycated hemoglobin under compression and shear exhibit an anisotropic mechanical behavior. *Journal of Biomolecular Structure and Dynamics*, *36*(6), 1417–1429.

Zou, X., Ma, W., Solov'Yov, I. A., Chipot, C., & Schulten, K. (2012). Recognition of methylated dna through methyl-cpg binding domain proteins. *Nucleic acids research*, *40*(6), 2747–2758.

Certificate of Attendance

American Physical Society
March Meeting 2022
March 14–18, 2022 | Chicago & Online

THIS IS TO CERTIFY THAT

Mr. Jhulan Powrel

attended the American Physical Society's March Meeting.
Pre-meeting tutorials and short courses were held on March 12 and 13.

APS physics MARCH MEETING 2022

Don Wise, Senior Meetings Registrar
March 18, 2022

Certificate of Paper Presentation

This is to certify that

Mr. Jhulan Powrel

has successfully presented the paper titled

Binding Free Energy in the R-State Dimer structure of Sickle and Normal Hemoglobin Protein

at the

National Conference on Advances in Atmospheric and Material Science

held on Nov. 6-7, 2022 at

the Department of Physics, Amrit Campus, Tribhuvan University, Kathmandu, Nepal.

Prof. Dr. Lok Narayan Jha
Chief Guest

Prof Dr. Leela Pradhan Joshi
Head, Department of Physics
Amrit Campus, TU

Dr. Lok Bahadur Baral
Campus Chief
Amrit Campus, TU

Organized by:

Department of Physics
Amrit Campus
Tribhuvan University
Nepal

Supported by:

International Science Programme
Uppsala University
Sweden

In collaboration with:

IEEE ELECTRON DEVICES SOCIETY

NSWIP

SALNet
South Asian Lightning Network

Certificate of Appreciation

THIS CERTIFICATE IS PRESENTED TO

MR. JHULAN POWREL

Tribhuvan University, Nepal

FOR YOUR PRESENTATION DURING THE ANPA CONFERENCE 2021

Nmalakar

DR. NABIN MALAKAR
CONVENER, ANPA CONFERENCE 2021

DR. JAGAN DEVKOTA
PRESIDENT, ANPA

Nepal Physical Society

Ghantaghar, Kathmandu

Certificate of Participation

This certificate is awarded to

Mr. Jhulan Powrel **from**

Butwal Multiple Campus, TU, Nepal, for his presentation entitled

Thermal and Transport Property of Mutated Hemoglobin Protein of Sickle Cell

On Scientific Session of 37th Annual Convention, February 6, 2021.

Prof. Dr. Narayan Prasad Chapagain
President
Nepal Physical Society

Dr. Sunil Babu Shrestha
Chief Guest
Vice – Chancellor, NAST

Prof. Dr. Leela Pradhan Joshi
Chair, Scientific Committee
37th Annual Convention, NPS

February 6, 2021

Nepal Physical Society

Ghantaghar, Kathmandu, Nepal

International Conference on Frontiers of Physics -2022 (ICFP-2022)

This Certificate of appreciation is awarded to

Jhulan Powrel

from CDP, Tribhuvan University, Nepal for his oral presentation on the **"Analysis of Vdw, Electrostatics, Hydrogen Bond, Salt Bridge Interaction and Steered Molecular Dynamics Study of Normal Hemoglobin Protein"** in the **International Conference on Frontiers of Physics -2022** held on January 22-24, 2022 via virtual platform.

Prof. Dr. Narayan P. Chapagain
Conference Chair, ICFP-2022
President, Nepal Physical Society

February 1, 2022

APPENDIX

A. Articles published in International Journals

Powrel, J., Koirala, R. P., & Adhikari, N. P. (2022). Structural stability of R-state conformation of carbonmonoxyl sickle and normal hemoglobin dimer. *Journal of Biomolecular Structure and Dynamics*, *Advance online publication.*(DOI: 10.1080/07391102.2022.2150890), 1-10. (PMID: 36473709)

Powrel, J., & Adhikari, N. P. (2022). Elastic property of sickle and normal hemoglobin protein: Molecular dynamics. *AIP Advances*, *12*(4), 045308.

Koirala, R. P., Thapa, B., Khanal, S. P., Powrel, J., Adhikari, R. P., & Adhikari, N. P. (2021). Binding of SARS-CoV-2/SARS-CoV spike protein with human ACE2 receptor. *Journal of Physics Communications*, *5*(3), 035010.

B. Articles Published in National Journals

Powrel, J., & Adhikari, N. P. (2021). Thermal properties of normal and sickled hemoglobin protein. *BIBECHANA*, *18*(1), 140–148.

Powrel, J. (2022). Computational estimation of free energy using fortran code. *Butwal Campus Journal*, *5*(1), 143–159. Retrieved from `https://nepjol.info/index.php/bcj/article/view/50199` doi: 10.3126/bcj.v5i1.50199

C. Articles in Review

- Elastic Property of Alpha Chain in Sickle and Normal Hemoglobin Protein by Steered Molecular Dynamics.
- Temperature Dependency of Transport Property of Dimer in Sickle and Normal Hemoglobin Protein.

D. Participation in Conferences, Seminar and Meeting

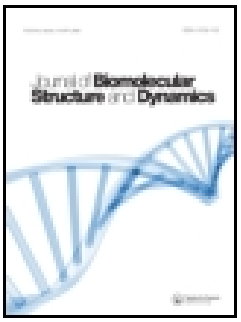
- "APS March Meeting 2022" organized by American Physical Society, USA, March 14-18, 2022.
  Oral presentation: "Analysis of Steered Molecular Dynamics Study of Sickle Hemoglobin Protein in Water"

- "National Conference on Advances in Atmospheric and Material Science" organized by Department of Physics, Amrit Science Campus, Thamel, Kathmandu, Nepal, Nov. 6-7, 2022.
  Oral presentation: "Binding Free Energy in the R-State Dimer Structure of Sickle and Normal Hemoglobin Protein"

- "ANPA conference" organized by Association of Nepali Physicists in America, July 9, 2021.
  Oral presentation: "Structure and SMD Study of Sickled Hemoglobin Protein"

- "International Conference on Frontiers of Physics" organized by Nepal Physical Society, Nepal, January 22-24, 2022.
  Oral presentation: "Analysis of vdW, Electrostatics, Hydrogen Bond, Salt Bridge Interaction and Steered Molecular Dynamics Study of Normal Hemoglobin Protein"

- "Scientific Session of 37th Annual Convention" organized by Nepal Physical Society, Feb. 6, 2021.
  Oral presentation: "Thermal and Transport Property of Mutated Hemoglobin Protein of Sickle Cell"



Journal of Biomolecular Structure and Dynamics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

Structural stability of R-state conformation of carbonmonoxyl sickle and normal hemoglobin dimer

Jhulan Powrel, Rajendra Prasad Koirala & Narayan Prasad Adhikari

To cite this article: Jhulan Powrel, Rajendra Prasad Koirala & Narayan Prasad Adhikari (2022): Structural stability of R-state conformation of carbonmonoxyl sickle and normal hemoglobin dimer, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2022.2150890

To link to this article: https://doi.org/10.1080/07391102.2022.2150890

Published online: 06 Dec 2022.

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
https://doi.org/10.1080/07391102.2022.2150890

Structural stability of R-state conformation of carbonmonoxyl sickle and normal hemoglobin dimer

Jhulan Powrel[a,b], Rajendra Prasad Koirala[b] and Narayan Prasad Adhikari[b]

[a]Department of Physics, Butwal Multiple Campus, Butwal, Nepal; [b]Central Department of Physics, Tribhuvan University, Kathmandu, Nepal

Communicated by Ramaswamy H. Sarma

ABSTRACT

A mutation at the sixth residue, glutamic acid to valine, in beta chain of hemoglobin distorts the entire shape of hemoglobin into a sickle shape. The investigation of the binding mechanisms of different chains of hemoglobin under the mutated condition can give an understanding of the molecular distortion. In this work, we have studied the binding mechanism between two chains in the dimer structure of the R-state conformation of carbonmonoxyl sickle hemoglobin and is compared with that of normal hemoglobin by using molecular dynamics simulations. The binding strength between α-chain (PROA) and β-chain (PROB) in hemoglobin dimer has been analyzed by estimating hydrogen bonds, salt bridges, hydrophobic interactions and non-bonded interactions (electrostatics and van der Waals). The quantitative estimation of aforementioned interactions depicts that the structural stability of normal hemoglobin dimer is found to be greater than that of sickle one. The outcomes of such interactions are also supported by the estimated free energy between the chains in R-state conformation of the dimers. The difference of binding free energy, calculated by utilizing the umbrella sampling technique, is found to be $\approx (0.67 \pm 0.06)$ kcal/mol.

ARTICLE HISTORY
Received 1 June 2022
Accepted 12 November 2022

KEYWORDS
Dimer; R-state conformation; hydrophobic interaction; hydrogen bond; free energy

Introduction

Hemoglobin is an iron-containing globular protein in the red blood cells of animals (Jia et al., 2016). Small bubbles of oxygen are combined with hemoglobin molecules in the lungs and thereby transported throughout the body to keep the tissues alive (Perutz, 1978; Krogh, 1941). These molecules provide the red color to the blood after the combination with oxygen (Billett, 1990; Antonini & Brunori, 1970). They move freely in the red blood cells. Besides oxygen transportation, hemoglobin has the other several functions like, genetic resistance of malaria, molecular heat transducer through its oxygenation–deoxygenation cycle and the enzymatic activities (Giardina et al., 1995). Hemoglobin can also interact with other three biologically important gases: carbon dioxide (CO_2), carbon monoxide (CO) and nitric oxide (NO) (Schechter, 2008).

The human hemoglobin molecule has an assembly of four globular protein subunits. Each such subunit is composed of a protein chain tightly coupled with a non-protein heme group. The four polypeptide chains are bound to each other by salt bridges, hydrogen bonds and hydrophobic interactions. There are two kinds of contacts between the α and β chains (Figure 1(a)) as α_1 β_1 and α_2 β_2 (Wu et al., 2008). A normal hemoglobin (Hb) molecule has the molecular weight 64,500 containing two α and two β polypeptide chains (Figure 1(b)) (Northrop & Anson, 1929). It is synthesized in multiple steps. The heme part containing an iron is synthesized in a series of steps in the mitochondria and the cytosole of immature red blood cells, while the globular protein parts are synthesized by ribosomes in the cytosole. In order to protect the heme part, globular protein surrounds it. Each single hemoglobin molecule has two globular chains, each with its own heme protein attached as shown in Figure 1(c). Two hemoglobin molecules combine to produce functional hemoglobin tetramer.

The mutation in the hemoglobin molecule may change the physical as well as biological properties of the entire cell (Thom et al., 2013). Sickle cell anemia is one such variety of mutation in which there is adenine substitution in the sixth codon of the beta gene (GAG-GTG), thereby encoding the valine instead of glutamic acid in sixth position of the beta chain of hemoglobin (Noguchi & Schechter, 1981; Shrestha & Karki, 2013). The sickled hemoglobin fails to deliver oxygen properly to the body, hence reducing the life time of the cell. Such abnormal hemoglobin causes red blood cells to become rigid and sticky so that they cannot transport oxygen properly (Yesudasan et al., 2019). It causes many types of disorders in the human body, like painful swelling, risk of stroke in the brain, increasing risk of pathogenic infections (Habara & Steinberg, 2016; Barpanda, 2013; Akinsheye & Klings, 2010). Of course, the sickle cell mutation in hemoglobin is evolved as an adaptation to protect against malaria. It is basically common among people whose ancestors originated from former and current malaria zones of the world; for example Africa, Asia, the Middle and Far East,

CONTACT Narayan Prasad Adhikari narayan.adhikari@cdp.tu.edu.np Central Department of Physics, Tribhuvan University, Kirtipur, Kathmandu, Nepal.

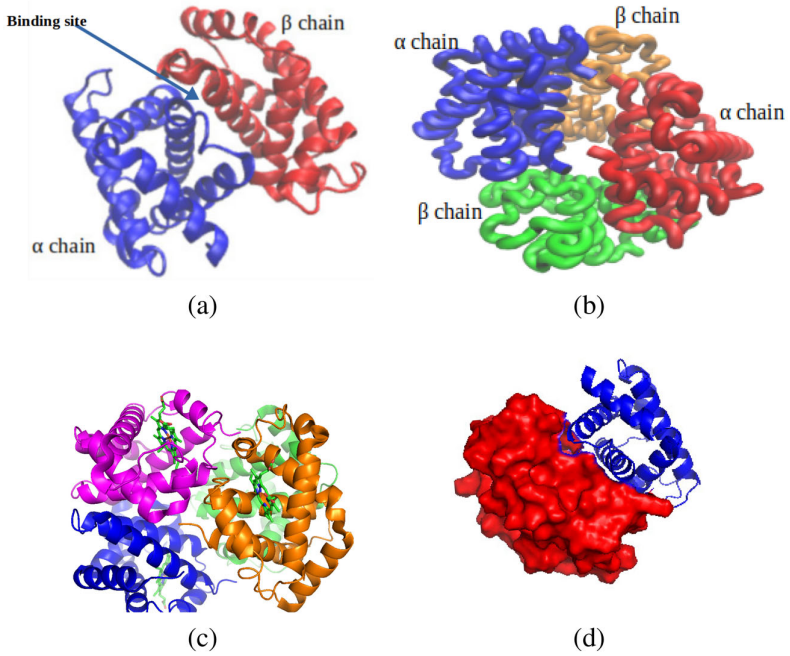


Figure 1. (a) Binding site of dimer in HbS (b) Binding site of chains A and B in the HbS dimer (c) Four heme legands in HbS (d) Surface interaction in $\alpha - \beta$ chain dimer of sickle hemoglobin.

Mediterranean islands and South America (Carter & Mendis, 2002; Weatherall, 2008).

The fundamental characteristics of sickle cell disease are studied from the deoxygenated (T-state) structure of sickled hemoglobin, which results in distorting the cell into a concavo-convex shape (Safo & Abraham, 2005; Ghatge et al., 2016). In this study we have taken R-state conformation of carbonmonoxyl sickle hemoglobin, which contains two chains representing the one α chain and another β chain. This structure contains the valine at sixth position of the β chain. To compare structural stability and atomic level interaction with normal form of dimer, the sixth position of such chain is mutated to glutamic acid as is present in normal hemoglobin (Ghatge et al., 2016; Perutz et al., 1968; Perutz & Lehmann, 1968). For the in-depth study of conformation and binding mechanism, we have estimated and compared binding free energy of the dimers at the interface (Figure 1(d)) for both conditions.

Methods and methodology

Molecular dynamics study has been carried out to investigate the intra- and inter-chain interactions in α and β chains of R-state conformation of carbonmonoxyl sickle normal hemoglobin dimer. Several interaction parameters like hydrogen bonding, non-bonded interactions (electrostatic and van der Waals), hydrophobic interaction and salt bridges have been estimated by applying nanoscale molecular dynamics (NAMD; Phillips et al., 2005) trajectory tools in visual molecular dynamics (VMD) (Humphrey et al., 1996). Umbrella sampling has been utilized to evaluate the trajectory for different samples, and the free energy was estimated by using the Weighted Histogram Analysis Method (WHAM) (Kumar et al., 1992).

Theoretical background

Nonbonded interactions

Electrostatics and van der Waals interactions play an important role in intramolecular as well as intermolecular binding in biomolecules. In the hemoglobin molecule, such interactions contribute significantly to binding PROB and PROA during the formation of a quaternary structure (Ghatge et al., 2016). The partial charge in atoms that forms dipole moment in the amino acids provides the electrostatic interaction, and the electron cloud shifting during the oscillation of atoms produces the vdW interaction. Such types of non-bonded interactions have significant contribution in intra- and intermolecular conformation (Fischer et al., 2002; DiStasio et al., 2014; Frenkel & Smit, 2002; González-Ruiz & Gohlke, 2006). Total interaction energy due to electrostatic and van der Waals interactions is given by (Lennard-Jones, 1931),

$$V_{\text{non-bonded}} = \frac{q_i q_j}{4\pi\varepsilon_m r_{ij}} + 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \quad (1)$$

where, ε_m denotes the permittivity of the medium, ϵ defines the interaction strength, σ defines the length scale, q_i and q_j are the two charges separated by distance r_{ij}.

Free energy

The study of a molecule during the decoupling of its different chains can reveal their unbinding mechanism, which in turn quantifies the binding strength of the entire molecule. The umbrella sampling method is considered a reliable

technique for the computational estimation of free energy. For this, different samples of configuration states of a molecule are taken in order to represents the entire phase space under a specific statistical ensemble. Then, these samplings are executed for the umbrella sampling MD simulation by applying the harmonic potential between two parts of a sample, taking appropriate force constant (Kästner, 2011). Free energy estimation is usually applied to study the binding mechanism. The binding free energy ΔG^0 of the molecule is estimated with (Raha & Merz Jr, 2005; Murcko, 1995),

$$\Delta G^0 = G^0(\text{complex}) - G^0(\text{alpha}) - G^0(\text{beta}) \quad (2)$$

where, $G^0(\text{complex})$ is the free energy of the complex of alpha-beta dimer, $G^0(\text{alpha})$ is the free energy of alpha chain and $G^0(\text{beta})$ is the free energy of beta chain in the complex. Twelve windows were prepared in sickle dimer and normal hemoglobin $\alpha-\beta$ dimer complexes. Every successive window was taken translating PROB 1 Å along the negative *x*-direction by taking PROA as the reference. The window size ensures the sufficient overlapping of successive windows to cover the reaction coordinate space. The reaction coordinate was chosen as the distance between the center of mass (COM) of α and β chain along the negative *x*-axis. To make the necessary overlapping reaction coordinates, a bias potential $V(\xi, i)$ was used to force the system to fluctuate in coordinate space, which is given by (Kästner, 2011),

$$V(\xi, i) = \frac{1}{2}k(\xi - \xi_i)^2 \quad (3)$$

where, i is an integer ranging from 0 to $n = 1$, k is the force constant, ξ is a collective variable along which the free energy is calculated, and ξ_i is the center of the biasing potential.

Molecular dynamics simulation

System setup

We have taken the structure of the R-state conformation of carbonmonoxyl sickle hemoglobin from PDB ID 5E6E. This structure contains the valine at its sixth position in PROB. Another structure to form normal hemoglobin was obtained by mutating from VAL6 to GLU6 by using the CHARMM-GUI (Lee et al., 2016) online web server. The new psf and pdb of both molecules were generated from the CHARMM-GUI web server by taking topology information from the modified CHARMM36 (i.e., CHARMM36m) force field (Huang et al., 2017). Then, each complex was solvated by TIP3P (Jorgensen et al., 1983) water in a cubical box of dimensions $(90 \times 90 \times 90)$ Å^3 under the periodic boundary condition (PBC). The sickle protein is solvated with 23818 (TIP3P) water molecules. The normal hemoglobin system is made with 11153 (TIP3P) water molecules. In order to electrically neutralize the system, Na^+ and Cl^- ions are added to yield a system of 0.15 m/L molar concentration.

Simulation details

Molecular dynamics (MD) simulations have been conducted by using the NAMD simulation package (Phillips et al., 2005). The parameter files are taken from the CHARMM36m force field (Huang et al., 2017). To remove the undesirable clashing of the atoms and unwanted effects of solvent in the system, the energy minimization (EM) run was executed for 10,000 steps by using conjugate gradient and line search algorithms. The EM run removes the possible artifacts during the construction of pdb structure, which may be ordered in low entropy and high regularity mode. Long-range interaction was addressed by using Particle Mesh Ewald (PME) method, keeping a 12 Å non-bonded cutoff. The NVT ensemble was then used to perform an equilibration run at 310 K with a time step of 2 fs per step. Finally, the production run was conducted for 100 ns under the NPT ensemble taking a time scale of 2 fs per step. The standard requirement of system size was maintained with at least five water molecules between any protein and the box boundary (Brooks et al., 1983). The velocity rescaling (El Hage et al., 2018) with 0.1 ps and the Parrinello Rahman (Bussi et al., 2009) methods were used for temperature and pressure control, respectively. This method of velocity rescaling is an extension of the Berendsen thermostat to which a stochastic force has been chosen so as to generate a correct canonical distribution (El Hage et al., 2018).

Umbrella sampling

An umbrella sampling technique was used to estimate the binding free energy of sickle and normal hemoglobin dimers. The purpose of umbrella sampling in MD simulation is to sample all the possible states in which a system of molecule of interest may exist. Based on this sampling idea, probability for the molecule to be in any of these energy states can be readily calculated. Certain states of the protein are separated from others by extremely high potential energy barriers. Hence, it takes a longer time to walk through all the molecular states. Umbrella sampling is the way to represent such phase space. We have utilized this technique to find the binding free energy of PROA and PROB of R-state sickle hemoglobin. To prepare the samples, PROB is displaced manually along the positive *x*-direction, taking PROA as the reference. The center of mass (COM) between these chains was taken as the reaction coordinate. The COM distance for the initial sample was taken at 23.8 Å. Then, the other samples were prepared by increasing the distances by 1 Å for each succeeding window (Figure 2). Thus, 12 windows were prepared for each of the systems. The final window was chosen in such a way that there was no significant interaction between the chains for both the systems.

For each window, energy minimization was carried out for 10,000 steps. Then, an equilibration run was executed at 10 ns for each under the NVT ensemble at 310 K and 1 atm pressure, taking 1 fs per step. Finally, the production run was propagated 20 ns for each sample, taking 2 fs per step under NPT conditions incorporating the appropriate collective variables.

Results and discussion

This work is basically focused on the study of structural conformation, the binding sites and the binding energy in

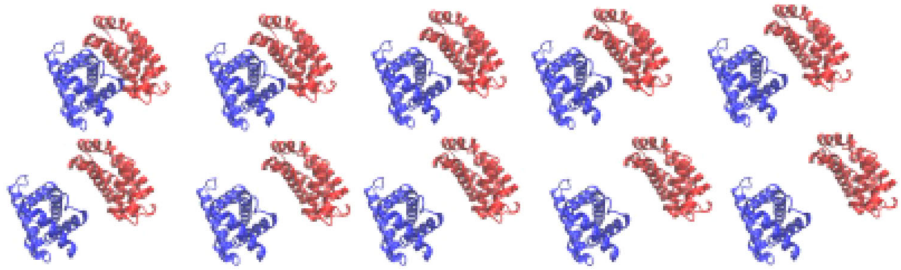


Figure 2. Snapshots of ten consecutive frames for umbrella sampling (US) of sickle hemoglobin dimer in which the separation increased by one Å in each consecutive window. The first window (upper left) to tenth window (lower left) are in clockwise order.

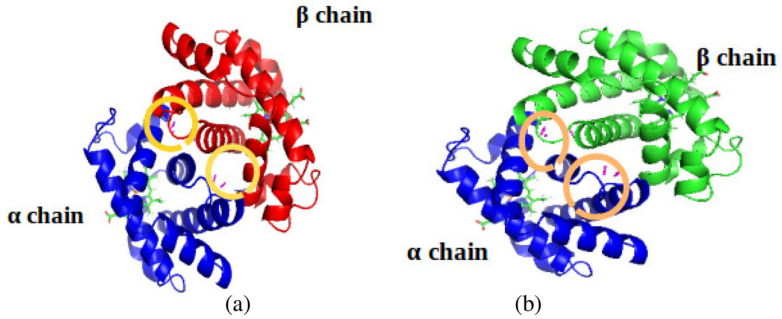


Figure 3. Polar binding sites (yellow circles) in (a) sickle hemoglobin dimer and (b) normal hemoglobin dimer.

between the two chains of the R-state conformation of carbonmonoxyl sickle hemoglobin. For this, molecular dynamics (MD) simulation has been carried out to understand the conformational stability. The results obtained from sickle hemoglobin has been compared to the R-state conformation of carbonmonoxyl normal hemoglobin. The binding free energy for these systems was estimated and compared in order to investigate the stability of complexes. The binding sites of sickle and normal hemoglobins are shown in Figure 3.

The PROA chain in the sickle hemoglobin dimer bound with the other atoms by seven polar bond in Figure 3(a). Figure 3(b) shows the seven polar bonds in normal hemoglobin, which bind the other atoms in the protein.

Root mean square deviation (RMSD)

RMSD has been utilized to investigate the structural stability of molecules in the aqueous environment. To examine the RMSD, we have taken the coordinates of the first frame of simulation as the reference. We have studied the RMSD of both normal and sickle proteins in order to study their structural stability. Figures 4 and 5 represent the RMSD plots of sickle and normal hemoglobin, respectively. A higher RMSD value indicating more flexibility of the protein and vice versa.

The RMSD plot of sickle dimer protein indicates that the structure becomes stable after 60 ns of simulation time. The PROB chain is achieving stability faster than the PROA chain and sickle dimer (Figure 4).

Figure 5 represents the RMSD plot of normal hemoglobin dimer. The graph indicates that in the normal hemoglobin protein PROA chain, PROB chain and whole dimer protein becomes stable faster than that of sickle protein. Figures 4 and 5 shows that the root mean square deviation (RMSD) of PROA is higher than of PROB and whole protein in both the dimer, that is, PROA chain is more flexible and the PROB chain is less flexible. The RMSD fluctuation of the sickle is greater. It indicates the unstable nature of the sickle dimer protein. The RMSD of each chain of two complexes and entire molecule have shown that the structures are stable in aqueous solution as shown in Figures 4 and 5. This shows that the structure is suitable for further study to evaluate the energetic and structural variations during the molecular dynamics simulations.

Hydrogen bonding

Hydrogen bonding is an important non-bonded interaction in the formation of single chains as well as in the interaction of different chains. This type of bonding contributes to both intra- and inter-molecular complex formation. Several biomedical activities are governed by hydrogen bonding between the atoms in the molecule.

In this study, we have investigated the pattern of intra-chain and inter-chain hydrogen bonding in each of the sickle and normal hemoglobin. In order to evaluate the hydrogen bonds (H-bonds), we used a cutoff distance of 3.5 Å and a cutoff angle of 30°. Figure 6 shows the variation of number of hydrogen bonds during 100 ns simulation. The number of intra-chain H-bonds in PROA and PROB is approximately the same; however the inter-chain H-bonds are significantly greater than each of the intra-chain.

The normal hemoglobin dimer, however, carries 60 inter-chain H-bonds on average as shown in Figure 7. It is seen that the normal hemoglobin dimer protein comprises a greater number of H-bonds than the sickle dimer. It is also obvious from the H-bond estimation that the normal hemoglobin dimer is more compact than a sickle dimer. We also

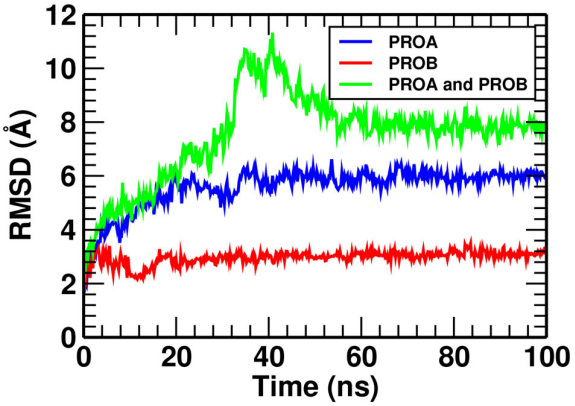


Figure 4. RMSD of sickle hemoglobin dimer.

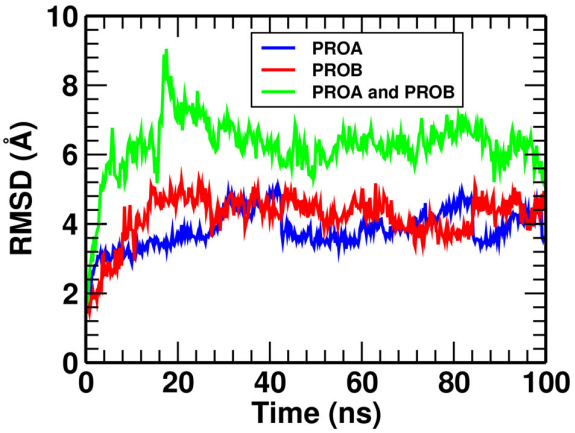


Figure 5. RMSD of normal hemoglobin dimer protein.

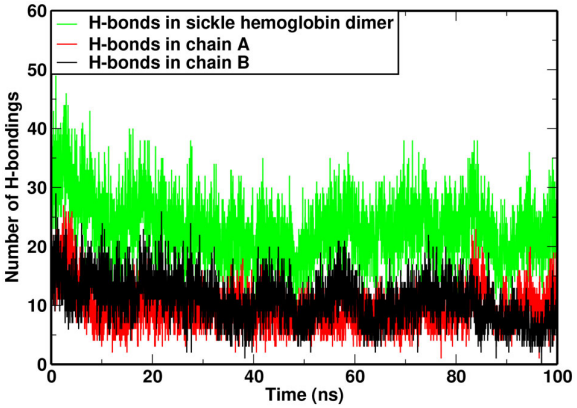


Figure 6. H-bond in sickle hemoglobin dimer.

estimated the H-bond occupancy in between PROA and PROB of both the dimers. In sickle dimer, the occupancy of H-bond between PROA and PROB, was 0.47, whereas the H-bond occupancy between PROA and PROB of normal hemoglobin was found to be 0.70. This result also supports the stronger binding conformation in normal dimer than that of sickle dimer. The H-bond occupancy for both the dimers above 40% are shown in Figures 8 and 9. In both conditions, the occupancy percentage in SER49-ASP47 was found the maximum and with value greater than 100%.

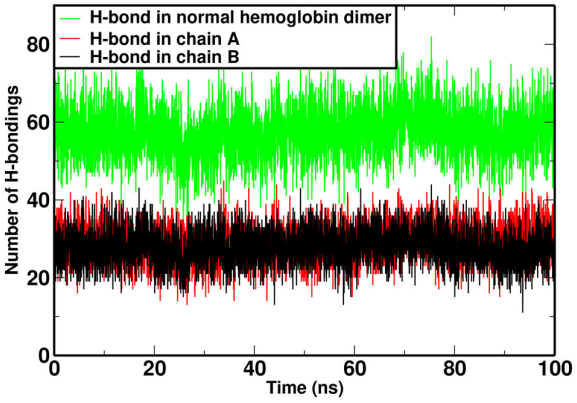


Figure 7. H-bond in normal hemoglobin dimer.

The H-bond occupancy in a normal hemoglobin dimer shows 19 residue pairs having occupancy above 40% (Figure 9). More residue pairs are involved in the H-bonding of normal hemoglobin dimer to provide the strength of elasticity in the protein.

Hydrophobic interaction

The non-bonded hydrophobic interaction provides the inter-residue interaction in the protein chain during the structural formation (Tina et al., 2007). To estimate the hydrophobic interaction, structure of both dimers were picked up taking five different time frames 500, 1500, 2500, 3500 and 4500 from the trajectory of simulation. From these frames, we have estimated the number of residue pairs involved in the intra-protein hydrophobic interactions of sickle and normal hemoglobin protein dimers (Table 1) using PIC.mbu software (Tina et al., 2007).

Hydrophobic interactions within the sickle and normal hemoglobin protein are studied separately. During simulation, more hydrophobic interaction residue pairs were found in the sickle dimer than in the normal hemoglobin dimer (Table 1). It indicates that hydrophobic interaction provides the greater compactness in normal hemoglobin than that of sickle hemoglobin protein.

Van der Waals (vdW) and electrostatic interaction

vdW and electrostatic interactions are the non-bonded interactions, which play important roles in the formation of protein structures. We have estimated such interactions in terms of non-bonded interaction with switch distances of 10 Å and cutoff distance of 12 Å. These interaction energies for sickle and normal hemoglobin are presented in Tables 2 and 3. We also carried out the quantitative estimation of the non-bonded binding interaction between protein, PROB chain and PROA chain in both dimers.

We also calculated the vdW and electrostatic interaction energy between the PROA and PROB chains of both dimers. The electrostatic interaction energy between PROA and PROB in a normal hemoglobin dimer is −67.58 kcal/mol, while it is −36.03 kcal/mol in a sickle dimer. Due to this, the sum of the vdW and electrostatic energy becomes much lower in a

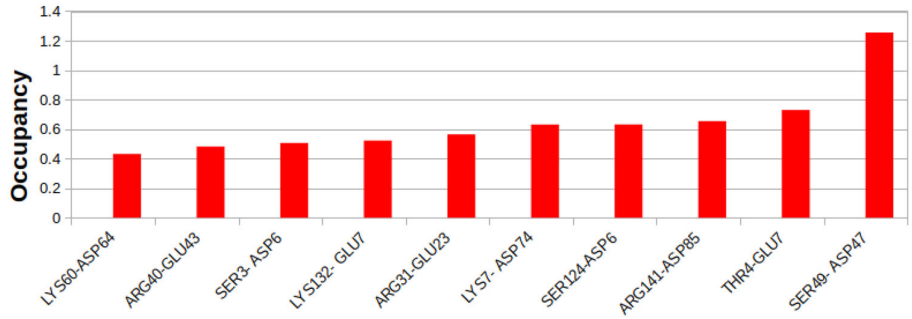


Figure 8. H-bond occupancy above 40% with residues involved in sickle hemoglobin dimer.

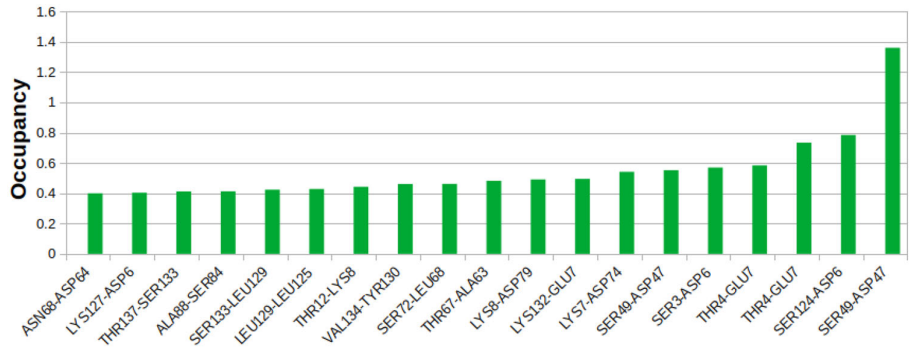


Figure 9. H-bond occupancy above 40% with residues involved in normal hemoglobin dimer.

Table 1. The number of residue pairs in the intraprotein hydrophobic interaction of sickle and normal hemoglobin dimers at five different time frames within the cutoff range of 5 Å.

Frame	500	1500	2500	3500	4500
Sickle Hb dimer	233	235	232	224	227
Normal Hb dimer	224	207	207	216	216

Table 2. vdW and electrostatic interaction energy (kcal/mol) in sickle Hb protein dimers.

For sickle Hb dimer	vdW	Electrostatic	Total (kcal/mol)
Protein	−884.62	−2719.82	−3604.45
PROA	−393.30	−1239.29	−1632.58
PROB	−457.01	−1443.64	−1900.65
Between PROA and PROB	−32.83	−36.03	−68.86

Table 3. vdW and electrostatic interaction energy (kcal/mol) in normal Hb protein dimers.

For normal Hb dimer	vdW	Electrostatic	Total (kcal/mol)
Protein	−924.05	−2852.37	−3776.41
PROA	−430.92	−1316.29	−1747.22
PROB	−462.27	−1467.60	−1929.87
Between PROA and PROB	−30.93	−67.58	−98.51

normal dimer (−98.51 kcal/mol) than in a sickle dimer (−68.86 kcal/mol). This shows stronger binding of PROA and PROB chains in normal hemoglobin than in sickle hemoglobin.

Salt bridges

Salt bridges are another important interaction that forms the stable structure of a protein. The occurrence of salt bridges in poly-peptide chains was found to be changed. Figures 10 and 11 represent the average of salt bridge occupancy in sickle and normal hemoglobin, respectively. We set the cut-off distance of 3.2 Å to analyze the salt bridges in the different chains in the proteins.

There are 32 pairs of active salt bridges in the sickle hemoglobin protein dimer, among them 15 pairs within PROA and 16 pairs within PROB (Figure 10). One inter-chain salt bridge is present between the residues of PROA and PROB in the sickle hemoglobin dimer. There are 51 pairs of active salt bridges in a normal hemoglobin protein dimer, among them 22 pairs in PROA, 20 pairs in PROB and the remaining 9 pairs of inter-chain salt bridges are in between PROA and PROB. ASP-LYS, GLU-LYS, GLU-ARG and ASP-ARG pairs are abundantly found in salt bridges of normal dimer (Figure 11). Mostly, the GLU90-LYS90 and ASP94-LYS90 of normal Hb have an occupancy above 50%, were found consistent during equilibration and have an almost equal occupancy. The occupancy for ASP74-LYS7, ASP6-LYS127, ASP79-LYS8 and ASP85-LYS139 is less than 4% (Figure 11).

Table 4. SASA of PROA chain, PROB chain and protein of both sickle and normal hemoglobin dimer in Å^2.

SASA of	PROA	PROB	Hb protein	Contact Area=$(S_1 + S_2 - S_3)/2$
	S_1 (Å)^2	S_2 (Å)^2	S_3 (Å)^2	S(Å)^2
Sickle Hb dimer	15433.12	13827.52	25979.74	1640.45
Normal Hb dimer	8635.63	9274.81	17496.69	206.88

Solvation accessible surface area (SASA)

Solvent accessible surface area is the evaluation of the region of the molecular surface exposed enough to the

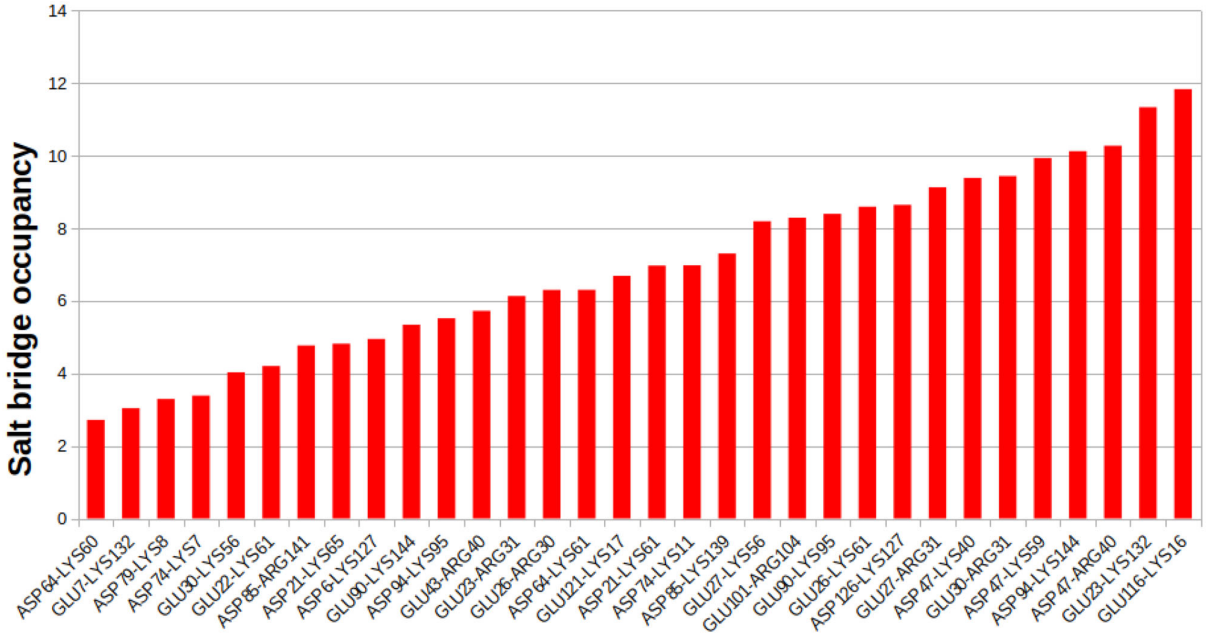


Figure 10. Salt bridge occupancy with residues involved in sickle hemoglobin dimer.

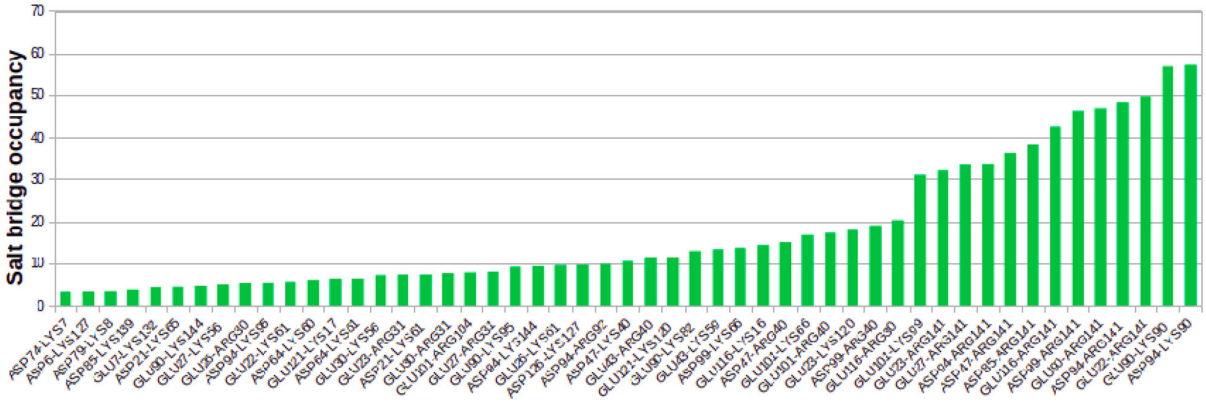


Figure 11. Salt bridge occupancy with residues involved in normal hemoglobin dimer.

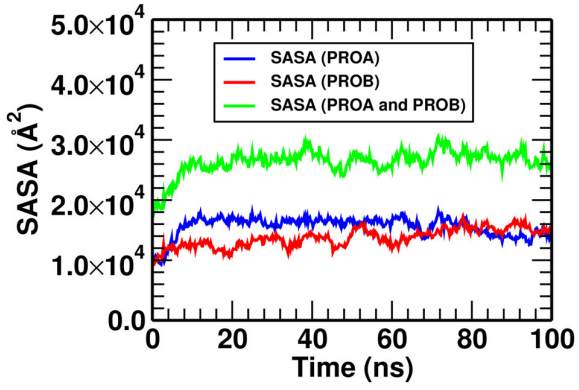


Figure 12. SASA's in sickle hemoglobin dimer.

solvent so as to interact with solvent molecules. SASA of sickle and normal hemoglobin are plotted as in Figures 12 and 13. Here, we have estimated the SASA of each chain (PROA and PROB) and that of the entire complex (Table 4). This procedure has been implemented for both the complexes. Then, the contact surface area of both chains of each system was determined by using the relation (Zou et al., 2012; Koirala et al., 2021a),

$$\text{Contact area(S)} = \frac{1}{2}[S_1(t) + S_2(t) - S_{12}(t)] \tag{4}$$

where, $S_1(t)$ is the SASA of the PROA and $S_2(t)$ is the SASA of PROB and $S_{12}(t)$ is the SASA in complex form (PROA and PROB). The contact surface area of two chains of a molecule is the surface at the interfacial region that contributes to binding the two chains. The greater contact surface area causes greater binding strength in between the chains. This assists in determining the stability of complex biomolecules.

It is a general fact that greater contact surface area means greater binding strength between chains. However, from this investigation of sickle and normal hemoglobin dimers, it is found that the contact surface area in the sickle is higher, whereas it has lower binding energy. This could be due to

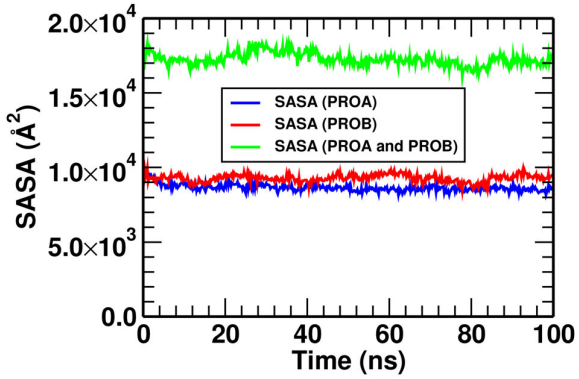


Figure 13. SASA's in normal hemoglobin dimer.

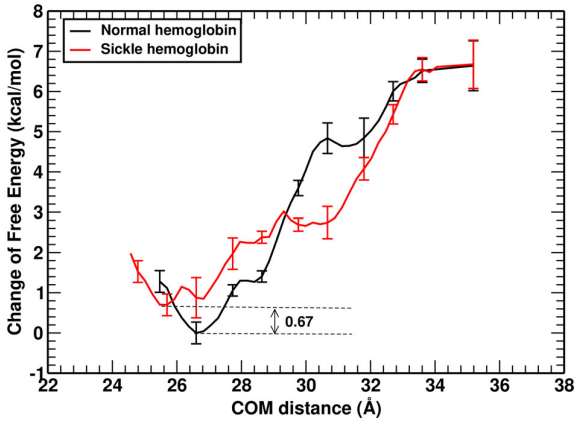


Figure 14. Free energy graph of sickle and normal hemoglobin dimer.

the fact that the sickle structure being curved in shape has a higher contact area than that of normal hemoglobin.

Binding free energy

The binding free energy between the chains in sickle and normal hemoglobin has been estimated by using the umbrella sampling method. For this, the beta chain (PROB) of each complex was translated along the positive × direction by taking PROA as the reference chain, and samples were prepared for each window taking 1 Å distance apart along the center of mass (COM) distance (Koirala et al., 2021b). Twelve samples were prepared for both the systems by examining the appropriate number of samples to estimate the binding free energy. Each of the samples is executed in a 20 ns simulation run under NPT conditions.

The free energy difference during the translation of the beta chain with respect to alpha chain is shown in Figure 14. From the figure, we see that the free energy difference for normal hemoglobin is higher than that for sickle hemoglobin, which shows the favourable binding in normal hemoglobin. Here, the free energy difference in sickle hemoglobin is ∼ 5.97 kcal/mol and that of normal hemoglobin is ∼ 6.64 kcal/mol. Hence, the free energy in normal hemoglobin is greater by (0.67 ± 0.06) kcal/mol than that of sickle hemoglobin. The error bars are taken from the standard deviation (SD) of data obtained from each window. We have changed scale in vertical axis to compare the free energy at the pitfall region of the plots. There is no overlap of error bars at the comparing region, which ensures the significant difference of free energy in sickle and normal hemoglobin dimer. In fact, the normal hemoglobin has higher values of interactions like vdW, electrostatic, salt bridge and hydrogen bond, which supports the quantitative estimation of binding free energy greater in normal dimer than that of sickle. During the interaction, we found that the contributions of vdW, electrostatic, salt bridge and hydrogen bond were found dominating effect in binding free energy. The greater number of residue pairs involved in the formation of hydrophobic interactions in sickle can cause greater contact area between the PROA and PROB of sickle than normal hemoglobin.

Conclusions and concluding remarks

The binding mechanism of the PROA and PROB chains of R-state carbonmonoxyl sickle hemoglobin has been studied by using molecular dynamics simulation method. The outcomes from sickle hemoglobin are compared with the results obtained from normal hemoglobin (mutation of valine-6 to glutamic-6 in the original pdb structure of 5E6E). The binding mechanism has been analyzed with different aspects like hydrogen bonding, non-bonded interaction (electrostatics and van der Waals), hydrophobic interaction, salt bridges etc. In this examination, the binding of two chains is stronger in normal hemoglobin (which contains glutamic acid in the sixth residue of the beta chain) than in sickle hemoglobin (which contains valine in place of glutamic acid). We then estimated the solvent accessible surface area (SASA) of both the complexes and contact areas in these conditions. The higher hydrophobic interaction in sickle dimer significantly governs the contact area of two chains in the dimer and hence the sickling. The mutated residue of the beta chain in the dimer and the nearest residue of the alpha chain are not so close to each other. Even the interfacial properties show a difference because of the differences in electronic interactions at the atomic level. The binding free energy for both the complexes has been estimated and compared. The free energy difference in normal hemoglobin is ∼ 6.64 kcal/mol and that of sickle hemoglobin is ∼ 5.97 kcal/mol, which shows a (0.67 ± 0.06) kcal/mol greater free energy difference in normal hemoglobin than that of sickle hemoglobin. It indicates that more energy is released from normal hemoglobin than that from sickle hemoglobin within the same change in reaction coordinates. Further study is awaited for the estimation of interfacial tension and thermal conduction.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

JP acknowledges University Grants Comission, Nepal for the Ph.D. Reasearch Grants of S & T-13/2076-77. We acknowledges the TWAS research grants RG 20-316.

Authors' contribution

JP carried out all the works, RPK supports for simulation run and NPA conceived the idea for the study. All authors read and approved the final manuscript. All data and paper are available with the correspondent author.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

References

Akinsheye, I., & Klings, E. S. (2010). Sickle cell anemia and vascular dysfunction: The nitric oxide connection. *Journal of Cellular Physiology*, *224*(3), 620–625. https://doi.org/10.1002/jcp.22195

Antonini, E., & Brunori, M. (1970). Hemoglobin. *Annual Review of Biochemistry*, *39*(1), 977–1042.

Barpanda, S. S. (2013). Use of image processing techniques to automatically diagnose sickle-cell anemia present in red blood cells smear. Unpublished doctoral dissertation, Department of Electrical Engineering, National Institute of Technology Rourkela.

Billett, H. H. (1990). *Hemoglobin and haematocrit* (3rd ed.). The History, Physical, and Laboratory Examinations.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. A., & Karplus, M. (1983). Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, *4*(2), 187–217. https://doi.org/10.1002/jcc.540040211

Bussi, G., Zykova-Timan, T., & Parrinello, M. (2009). Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *The Journal of Chemical Physics*, *130*(7), 074101. https://doi.org/10.1063/1.3073889

Carter, R., & Mendis, K. N. (2002). Evolutionary and historical aspects of the burden of malaria. *Clinical Microbiology Reviews*, *15*(4), 564–594. https://doi.org/10.1128/CMR.15.4.564-594.2002

DiStasio, R. A., Gobre, V. V., & Tkatchenko, A. (2014). Many-body van der Waals interactions in molecules and condensed matter. *Journal of Physics: Condensed Matter*, *26*(21), 213202.

El Hage, K., Hedin, F., Gupta, P. K., Meuwly, M., & Karplus, M. (2018). Valid molecular dynamics simulations of human hemoglobin require a surprisingly large box size. *eLife*, *7*, e35560. https://doi.org/10.7554/eLife.35560

Fischer, W.-N., Loo, D. D. F., Koch, W., Ludewig, U., Boorer, K. J., Tegeder, M., Rentsch, D., Wright, E. M., & Frommer, W. B. (2002). Low and high affinity amino acid h+-cotransporters for cellular import of neutral and charged amino acids. *The Plant Journal : For Cell and Molecular Biology*, *29*(6), 717–731. https://doi.org/10.1046/j.1365-313x.2002.01248.x

Frenkel, D., & Smit, B. (2002). *Understanding molecular simulation*. Academic Press.

Ghatge, M. S., Ahmed, M. H., Omar, A. S. M., Pagare, P. P., Rosef, S., Kellogg, G. E., Abdulmalik, O., & Safo, M. K. (2016). Crystal structure of carbonmonoxy sickle hemoglobin in r-state conformation. *Journal of Structural Biology*, *194*(3), 446–450. https://doi.org/10.1016/j.jsb.2016.04.003

Giardina, B., Messana, I., Scatena, R., & Castagnola, M. (1995). The multiple functions of hemoglobin. *Critical Reviews in Biochemistry and Molecular Biology*, *30*(3), 165–196. https://doi.org/10.3109/10409239509085142

González-Ruiz, D., & Gohlke, H. (2006). Targeting protein-protein interactions with small molecules: Challenges and perspectives for computational binding epitope detection and ligand finding. *Current Medicinal Chemistry*, *13*(22), 2607–2625. https://doi.org/10.2174/092986706778201530

Habara, A., & Steinberg, M. H. (2016). Minireview: Genetic basis of heterogeneity and severity in sickle cell disease. *Experimental Biology and Medicine (Maywood, N.J.)*, *241*(7), 689–696. https://doi.org/10.1177/1535370216636726

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., Grubmüller, H., & MacKerell, A. D. (2017). Charmm36m: An improved force field for folded and intrinsically disordered proteins. *Nature Methods*, *14*(1), 71–73. https://doi.org/10.1038/nmeth.4067

Humphrey, W., Dalke, A., & Schulten, K. (1996). Vmd: viSual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Jia, Y., Duan, L., & Li, J. (2016). Hemoglobin-based nanoarchitectonic assemblies as oxygen carriers. *Advanced Materials (Deerfield Beach, Fla.)*, *28*(6), 1312–1318. https://doi.org/10.1002/adma.201502581

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935. https://doi.org/10.1063/1.445869

Kästner, J. (2011). Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(6), 932–942. https://doi.org/10.1002/wcms.66

Koirala, R. P., Pokhrel, R., Baral, P., Tiwari, P. B., Chapagain, P. P., & Adhikari, N. P. (2021b). Structural insights into the repair mechanism of agt for methyl-induced dna damage. *Biological Chemistry*, *402*(10), 1203–1211. https://doi.org/10.1515/hsz-2021-0198

Koirala, R. P., Thapa, B., Khanal, S. P., Powrel, J., Adhikari, R. P., & Adhikari, N. P. (2021a). Binding of sars-cov-2/sars-cov spike protein with human ace2 receptor. *Journal of Physics Communications*, *5*(3), 035010. https://doi.org/10.1088/2399-6528/abea27

Krogh, A. (1941). The comparative physiology of respiratory mechanisms. In *The comparative physiology of respiratory mechanisms*. University of Pennsylvania press.

Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. the method. *Journal of Computational Chemistry*, *13*(8), 1011–1021. https://doi.org/10.1002/jcc.540130812

Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., MacKerell, A. D., Klauda, J. B., & Im, W. (2016). Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Journal of Chemical Theory and Computation*, *12*(1), 405–413. https://doi.org/10.1021/acs.jctc.5b00935

Lennard-Jones, J. E. (1931). Cohesion. *Proceedings of the Physical Society*, *43*(5), 461–482. https://doi.org/10.1088/0959-5309/43/5/301

Murcko, M. A. (1995). Computational methods to predict binding free energy in ligand-receptor complexes. *Journal of Medicinal Chemistry*, *38*(26), 4953–4967. https://doi.org/10.1021/jm00026a001

Noguchi, C. T., & Schechter, A. N. (1981). The intracellular polymerization of sickle hemoglobin and its relevance to sickle cell disease. *Blood*, *58*(6), 1057–1068.

Northrop, J. H., & Anson, M. (1929). A method for the determination of diffusion constants and the calculation of the radius and weight of the hemoglobin molecule. *The Journal of General Physiology*, *12*(4), 543–554. https://doi.org/10.1085/jgp.12.4.543

Perutz, M. F. (1978). Hemoglobin structure and respiratory transport. *Scientific American*, *239*(6), 92–125. https://doi.org/10.1038/scientificamerican1278-92

Perutz, M. F., & Lehmann, H. (1968). Molecular pathology of human haemoglobin. *Nature*, *219*(5157), 902–909. https://doi.org/10.1038/219902a0

Perutz, M., Muirhead, H., Cox, J., & Goaman, L. (1968). Three-dimensional Fourier synthesis of horse oxyhaemoglobin at 2.8 å resolution: The atomic model. *Nature*, *219*(5150), 131–139. https://doi.org/10.1038/219131a0

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., & Schulten, K. (2005). Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, *26*(16), 1781–1802. https://doi.org/10.1002/jcc.20289

Raha, K., & Merz Jr, K. M. (2005). Calculating binding free energy in protein–ligand interaction. *Annual Reports in Computational Chemistry*, *1*, 113–130.

Safo, M., & Abraham, D. (2005). The enigma of the liganded hemoglobin end state: A novel quaternary structure of human carbonmonoxy

hemoglobin. *Biochemistry*, *44*(23), 8347–8359. https://doi.org/10.1021/bi050412q

Schechter, A. N. (2008). Hemoglobin research and the origins of molecular medicine. *Blood*, *112*(10), 3927–3938. https://doi.org/10.1182/blood-2008-04-078188

Shrestha, A., & Karki, S. (2013). Analysis of sickle hemoglobin. *Journal of Pathology of Nepal*, *3*(6), 437–440. https://doi.org/10.3126/jpn.v3i6.8989

Thom, C. S., Dickson, C. F., Gell, D. A., & Weiss, M. J. (2013). Hemoglobin variants: Biochemical properties and clinical correlates. *Cold Spring Harbor Perspectives in Medicine*, *3*(3), a011858. https://doi.org/10.1101/cshperspect.a011858

Tina, K., Bhadra, R., & Srinivasan, N. (2007). Pic: Protein interactions calculator. *Nucleic Acids Research*, *35*(Web Server issue), W473–W476.

Weatherall, D. (2008). Genetic variation and susceptibility to infection: The red cell and malaria. *British Journal of Haematology*, *141*(3), 276–286. https://doi.org/10.1111/j.1365-2141.2008.07085.x

Wu, T., Wang, X., Ge, H., & Cohen, B. (2008). Multi-scale and multi-physics modeling of sickle-cell disease - Part I molecular dynamics simulation, 11.

Yesudasan, S., Douglas, S. A., Platt, M. O., Wang, X., & Averett, R. D. (2019). Molecular insights into the irreversible mechanical behavior of sickle hemoglobin. *Journal of Biomolecular Structure & Dynamics*, *37*(5), 1270–1281. https://doi.org/10.1080/07391102.2018.1456362

Zou, X., Ma, W., Solov'yov, I. A., Chipot, C., & Schulten, K. (2012). Recognition of methylated dna through methyl-cpg binding domain proteins. *Nucleic Acids Research*, *40*(6), 2747–2758. https://doi.org/10.1093/nar/gkr1057

Elastic property of sickle and normal hemoglobin protein: Molecular dynamics

Cite as: AIP Advances **12**, 045308 (2022); https://doi.org/10.1063/5.0086539
Submitted: 26 January 2022 • Accepted: 18 March 2022 • Published Online: 06 April 2022

Jhulan Powrel and Narayan Prasad Adhikari

Elastic property of sickle and normal hemoglobin protein: Molecular dynamics

Cite as: AIP Advances 12, 045308 (2022); doi: 10.1063/5.0086539
Submitted: 26 January 2022 • Accepted: 18 March 2022 •
Published Online: 6 April 2022

View Online Export Citation CrossMark

Jhulan Powrel[1,2] and Narayan Prasad Adhikari[2,a)]

AFFILIATIONS

[1] Department of Physics, Butwal Multiple Campus, Butwal, Rupandehi, Nepal
[2] Central Department of Physics, Tribhuvan University, Kirtipur, Kathmandu, Nepal

[a)] Author to whom correspondence should be addressed: narayan.adhikari@cdp.tu.edu.np

ABSTRACT

This work focuses on identifying the conformational stability and binding components in sickle and normal hemoglobin to explore the elastic properties and realize the stiffness by using molecular dynamics simulation. Our investigation shows that a larger force is required to separate the beta chain of normal hemoglobin in comparison to the sickle hemoglobin by using steered molecular dynamic. In sickle hemoglobin protein (HbS), the hydrogen bond binding force of the beta chain is 7073.74–12 646.80 pN for pulling velocities of 0.000 20–0.000 40 nm/ps with the spring constant of 800 kcal mol^{-1} nm^{-2}. Similarly, in normal hemoglobin protein, the hydrogen bond binding force in the beta chain ranges from 12 005.00 to 17 753.70 pN for the same values of pulling velocities and spring constant. This indicates that the normal hemoglobin is stiffer than sickle hemoglobin. We have also analyzed the solvent accessible surface area (SASA) of both proteins, and our investigation shows that the SASA of normal hemoglobin is much less than that of sickle hemoglobin because of the sickled structure of HbS. We have also studied the van der Waals (vdW), electrostatic, hydrophobic, and salt bridge interactions in both kinds of hemoglobin. The sum of vdW, electrostatics, and hydrophobic interactions in HbS is higher, whereas salt bridge interactions are found lower in sickle normal hemoglobin proteins than in normal hemoglobin protein.

I. INTRODUCTION

A single hemoglobin molecule has two types of globin chains, each with its own heme protein. One globin chain is alpha and the other is beta. Two hemoglobin molecules combine to produce a functional hemoglobin tetramer (Fig. 1). In inherited hemoglobin, the amino acid sequence is altered because of incorrect deoxyribonucleic acid (DNA) code,[1] and such a cell having an abnormality is called a sickle cell. This abnormality causes sickle cell disease due to a change in the genetic code of hemoglobin protein in cells. Such a mutated hemoglobin protein (HbS) cannot carry oxygenated blood to the body in a regular manner. Sickle cell disease is also a molecular disease that was investigated over the last half-century.[2] A hemoglobin protein has four chains referred to as A (α), B (β), C (α), and D (β) (Fig. 1). Such an α chain of protein carries 141 amino acid residues and the β chain has 146 amino acid residues, i.e., a hemoglobin protein has a tetrameric structure with four globular chains. Sickle hemoglobin has a valine residue in its sixth position of the beta chain [Fig. 2(a)], whereas a normal hemoglobin protein comprises glutamic acid residue in the sixth position[3] of the beta (β) chain [Fig. 2(b)]. This structural abnormality in sickle hemoglobin [Fig. 2(a)] was pointed out as the replacement of the sixth glutamic acid residue in normal hemoglobin [Fig. 2(b)] protein by valine.[4] The deformation and stiffness of the hemoglobin cell are primarily caused in sickle cell disease, and it is important to differentiate sickle cell anemia from the anemias caused by other abnormal hemoglobin.[5]

Several experimental works have already been carried out to understand the genetic diseases due to the mutation in hemoglobin. On the other hand, heterozygotes for the sickle gene are relatively protected against the danger of dying of malaria, as now firmly established through numerous clinical field studies from different parts of Africa and Asia.[6] The mutation in hemoglobin occurs due to a change in nucleotide adenine (A) into nucleotide thymine (T), which forms abnormal beta-globin, generating deformed sickle cells.[7]

This mutated hemoglobin causes red blood cells to become rigid and sticky so that they cannot transport oxygen properly.

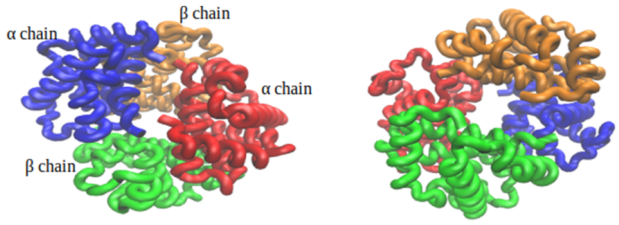


FIG. 1. Tetrameric structure of HbS molecule in globular form with front and back views.

Such an infected cell by sickle cell disease has a short lifetime. It causes many types of disorders in the human body. Computational simulation results show a significant variation in the mechanical strength in different directions (anisotropy) as well as a strong variation among the four different types of hemoglobin configurations. The glycated hemoglobin molecule has an overall higher compressive mechanical stiffness and shear stiffness when compared to the deoxyhemoglobin, oxyhemoglobin, and carboxyhemoglobin molecules.[8] Even though many studies have been done previously, the atomic in depth of the molecular configuration and elastic behavior have not been studied yet. As the mechanical strength of the hemoglobin is significantly changed after mutation in its sixth residue (glutamic acid to valine), the comparison of normal and sickle hemoglobin is very important to understand the molecular mechanism of hemoglobin protein[9] after mutation. Motivated by this significant change in mechanical properties in sickle hemoglobin, we have investigated the molecular configuration and elastic behavior of the molecules. The average values of Root Mean Square Deviation (RMSD) of atomic motions were observed at about 0.7 Å for HbA and 1.0 Å for HbS for a short simulation run.[10] The change in flexibility of the backbone atom may play an important role in HbS polymerization. Also, the atomic force microscopy (AFM) method is already implemented for estimating the Young modulus, and the results have been compared with the pathological sickle hemoglobin.[11] Many simulation works have been performed by molecular dynamics with different box sizes as well as longer simulation time scales for a better fit of the results with experimental results.[11] The bio-physical processes in both sickle and normal hemoglobin (1A3N) are not studied for further understanding of their elastic properties; hence, the present work is focused on relating the structural formation and their bio-physical process by using molecular dynamics simulation. We have performed the MD simulation of 100 ns for each sickle and normal hemoglobin. The time analysis of conformational energy, hydrogen bonding, and hydrophobic bonding contributes to the stability of the protein associated with both kinds of hemoglobin.

In Sec. II, we discuss the preparation of the system with the theoretical background. Sections III and IV include the discussion of results and the conclusions, respectively.

II. METHODS AND METHODOLOGY

A. System setup

To study the molecular interactions and the elastic properties of normal and sickle hemoglobin, we have taken two molecular structures, with PDB IDs: 1A3N[12] and 2HbS,[13] respectively. The normal and sickle hemoglobin have similar structures except the mutation at the sixth residue of the beta chain. The normal hemoglobin contains glutamic acid at its sixth residue in the beta chain, whereas it was mutated to valine in sickle hemoglobin. CHARMM-GUI[14] web server was used to create the protein structure (psf) files accompanied with pdb files. Then, the structure files were solvated by TIP3P[15] water box of dimensions (100 ×100× 500) $Å^3$ under the periodic boundary condition (PBC) in both molecular systems. There are 69 132 atoms in 1A3N solution and 53 427 atoms in 2HbS solution with enough Na^+ and Cl^- ions to yield the system of 0.15 m/l molar concentration.

B. Molecular dynamics simulations

This work has been carried out by the NAnoscale Molecular Dynamics (NAMD)[16] package. All-atom molecular dynamics simulations have been executed by using modified CHARMM36 force field[17] and TIP3P water model using 5.1.4 version of NAMD package on GPUS. Energy minimization run was performed for 10 000 steps. This simulation run has been executed to remove the unwanted steric hindrance of the atoms in the molecules.[18,19] Then, the temperature and pressure are maintained at the desired value by equilibration run taking 2 fs/step under the NVT ensemble. Finally, the production run was conducted for 100 ns under NPT ensemble taking 2 fs/step at temperature 310 K under atmospheric pressure. The standard requirement of at least five water molecules between any protein and the box boundary[20] was maintained for accuracy. For the long-range electrostatic interactions, Particle-Mesh Ewald (PME)[16] was used with a grid spacing of 1 Å and a cutoff of 10 Å together with a 10 Å of switching distance for the Lennard-Jones (LJ) interactions.[21] The LINCS algorithm[22] was used for constraining the hydrogen bonds (H-bonds) involving hydrogen atoms.[23] The velocity rescaling[22] with 0.1 ps and Parrinello–Rahman[24] method were used for temperature and pressure control, respectively.

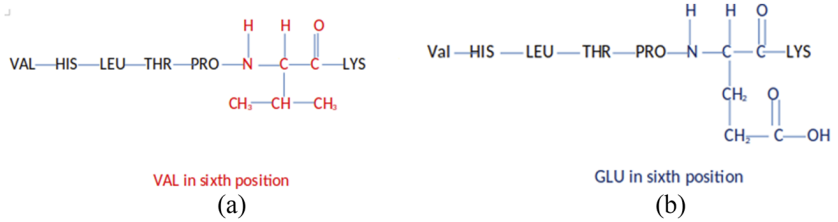


FIG. 2. (a) Valine residue of sickle hemoglobin in the sixth position of the beta chain. (b) Glutamic acid residue of normal hemoglobin in the sixth position of the beta chain.

C. Non-bonded interactions

Electrostatic and van der Waals interactions play an important role in intramolecular as well as in intermolecular binding in hemoglobin protein. Each of the β and α chains of both hemoglobin proteins possesses a large number of peptide molecules containing the net negative charge. Likewise, some of the amino acids have non-zero, positive or negative charges. In addition to them, atoms in the molecules have non-zero partial charges. In many situations, they form the dipole moments. These partial or net charges in the molecule play a significant role in binding the chains in protein molecules via electrostatic interaction.[25] On the other hand, van der Waals interaction is special in all types of atoms and assists in strengthening the binding of the molecules in the complex.[26] The non-bonded interaction consists of electrostatic[27] and van der Waals interactions,[28] which is given by

$$V_{non-bonded} = \frac{Qq}{4\pi\varepsilon_m r} + 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right], \quad (1)$$

where ε_m denotes the permittivity of the medium, ϵ defines the interaction strength, σ defines the length scale, Q and q are the two charges separated by distance r, and r_{ij} is the distance between two atoms i and j interacting by vdW.

D. Steered molecular dynamics (SMD)

Steered molecular dynamics (SMD) is often used to study the unfolding pathways of a globular protein. This technique is also employed to understand the elastic behavior of the bio-molecules. In particular, the introduction of hydrogen bonds into the backbone and side-chain as well as the incorporation of the unnatural sixth valine in fibrous amino acids with a beta chain have greatly expanded the ability of SMD to obtain the specific structural and dynamic information.[29] In this work, we have utilized this method to estimate the force to decouple the hydrogen bonds. SMD allows the exploration of biological processes on a time scale accessible to molecular dynamics simulations.[30] In actual SMD simulation, water plays a fundamental role in unfolding pathways for proteins. In the SMD simulations at constant velocity protocol, the force experienced by the pulled terminal residue is given by

$$F = k(vt - x), \quad (2)$$

where k is the spring constant, v is the pulling velocity, and x is the displacement of the SMD atom. During the simulations, we have fixed one atom of the beta chain and pulled the atom of another end to stretch and unfold the protein.[16] Before running the actual simulations, we have prepared the system in water. The stretching of the beta chain is achieved through the constant pulling velocity offered by the N–C terminal SMD method (Fig. 3). N–C terminal SMD explores the elastic property of materials. It also explores the pathways of the unfolding of protein and shows the occurrence of biological processes on time scale. In this type of simulation, the SMD atom is attached to a dummy atom via a virtual spring. This dummy atom is moved at constant velocity and then the force between both is measured using the following equation:[31,32]

$$\vec{F} = -\vec{\nabla} U, \quad (3)$$

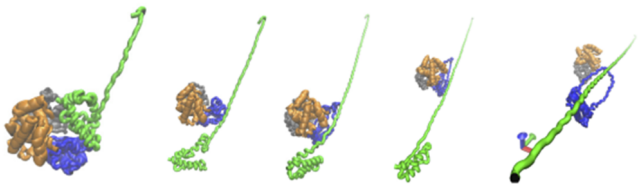


FIG. 3. Snapshots of five consecutive frames of N–C terminal SMD of the beta chain of sickle hemoglobin.

where the potential energy (U) is

$$U = \frac{1}{2}k\left[vt - \left(\vec{r} - \vec{r_0}\right)\cdot\hat{n}\right]^2, \quad (4)$$

with r and r_0 being the final and the original positions of the SMD atom, respectively, and $\hat{n}$ being the direction of pulling.

Pulling velocities and spring constant are chosen in such a way that their effects do not diminish during the simulation. The virtual spring between the dummy atom and the SMD atom has a spring constant k in the range of 500–1100 kcal mol^{-1} nm^{-2}. Therefore, the pulling velocities have been considered within the range of 0.0002–0.0004 nm/ps and the spring constant k within the range of 500–1100 kcal mol^{-1} nm^{-2}.

III. RESULTS AND DISCUSSION

Molecular dynamics (MD) simulation has been carried out to understand the conformational stability and the elastic behavior of normal and sickle hemoglobin. The properties are compared between these molecules. In this section, we discuss the findings of this work. First, the Root Mean Square Deviation (RMSD) of sickle and normal (de-oxy human hemoglobin) proteins is compared. Then, we present the different non-bonded interactions. Finally, we discuss the SMD simulation.

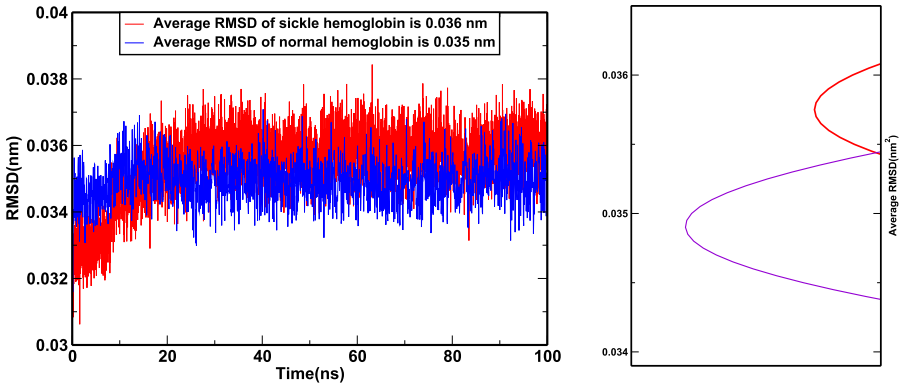
A. RMSD

RMSD has been utilized to investigate the structural stability of the molecules in the aqueous environment. To examine the RMSD, we have taken the coordinates of the first frame of simulation as the reference. We have studied the RMSD of both normal and sickle proteins in order to study their structural stability. Figure 4 represents the RMSD plot of HbS and A3N.

As shown in Fig. 4(a), the RMSD is found stable after 20 and 25 ns simulation time scale in HbS and A3N, respectively. The average values of RMSD are 0.036 and 0.035 nm for HbS and A3N, respectively [Fig. 4(b)], which indicated that both are stable structures.

B. Hydrogen bond (H-bond) interaction

Hydrogen bonding is an important interaction in the formation of single chain as well as in the interaction of different chains in the quaternary structure of a protein. This type of bonding system is not only involved in the formation of single protein structure but also in the formation of complex structure between two or more proteins or ligands. Several biological characteristics of the molecule depend on the structure of the molecule, which can be governed by



(a) (b)

FIG. 4. (a) RMSD of sickle and normal hemoglobin at a temperature of 310 K with time scale. (b) Average RMSD of HbS (red) and A3N (blue).

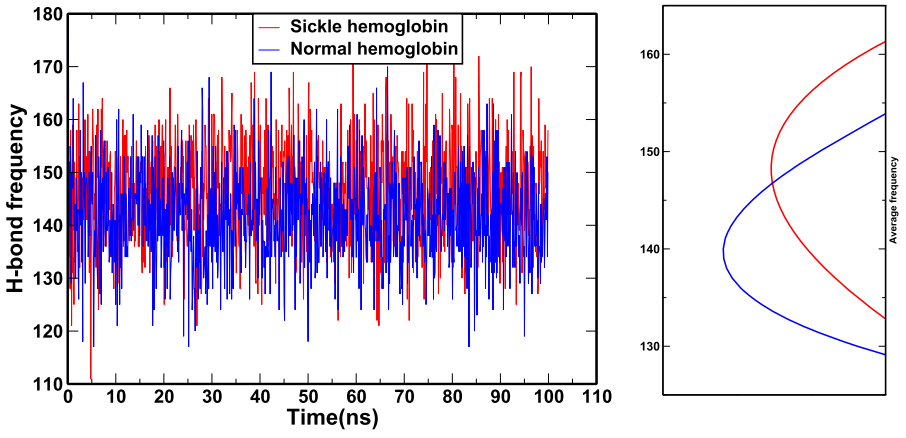
the pattern of hydrogen bonding. Hydrogen bonds are responsible for determining the three dimensional structure of folded proteins and play an important role in their bio-chemical changes. To study the residues involved in the hydrogen bonding and to estimate the number of hydrogen bonds between α and β chains of HbS and A3N, we used the cutoff parameters of 0.32 nm and 30° for length and angle, respectively. Figure 5 shows the variation occurring in the formation of a number of hydrogen bonds in between the residues under study.

The average number of hydrogen bond formations in HbS and A3N is found to be ≈147 and ≈142, respectively, as shown in Fig. 5(a). In Fig. 5(b), it is noted that the sickled hemoglobin has a greater average number of hydrogen bond interactions than that of normal hemoglobin.

The conformational stability of normal and sickle hemoglobin has been analyzed by estimating the hydrogen bond occupancy percentage. For this, we have taken the occupancy percentage of both proteins above 100%. Figure 6 represents the occupancy of hydrogen bonds above 100% in the sickled (red) and normal hemoglobin (black). The results showed that there were 428 pairs of residues involved in hydrogen bonds of HbS with different occupancies. Among them, 29 pairs have occupancy above 100 (Fig. 6). The most stable H-bonding is of SER49-ASP47 with occupancy of 310 (Fig. 6). In normal hemoglobin, there were 378 pairs of residues involved in H-bonding with various occupancies. Among them, 32 pairs have shown occupancy above 100. The maximum occupancy of H-bond is obtained in SER49-ASP47 residue with occupancy of 298. LYS residues are most abundantly found in H-bond, and SER49-ASP47 residue pair is consistent with H-bonding throughout the simulation in both proteins. The most stable hydrogen bonding is found to be of SER49-ASP47 residue with occupancy of 310. This study estimated that the hydrogen bonds are stronger in HbS.

C. Hydrophobic interaction

Hydrophobic interaction reveals the inter residue interaction of protein during the formation of its structure.[33] The hydrophobic nature of surface residues minimizes the overall size of the molecule so that the inter residue interaction enhances and the nature is reversed in the molecules containing hydrophilic residues on the surface. Figures 7(a) and 7(b) show the hydrophobic and hydrophilic surfaces of valine and glutamic acid residue of the beta chain in sickle and normal hemoglobin, respectively.



(a) (b)

FIG. 5. (a) H-bond frequency of occurrence in HbS and A3N. (b) Average number of H-bonds in HbS (red) and A3N (blue).

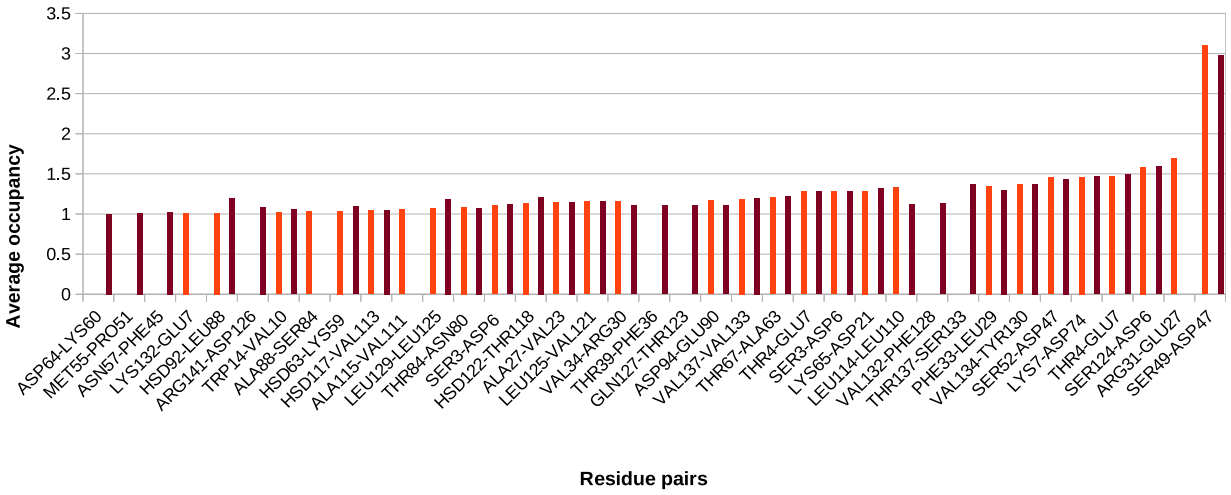


FIG. 6. Graph of occupancy percentage of H-bond vs residue pairs in sickle and normal hemoglobin.

Hydrophobic interactions within the normal and sickle hemoglobin are studied separately. During simulation, 491 and 303 hydrophobic pairs were found in sickled and normal hemoglobin, respectively. Then, we further analyzed the inter-chain hydrophobic interactions within the molecules. In sickle hemoglobin, 32 pairs of hydrophobic residues are active in between protein–protein chains; out of them, 16 pairs are with only the beta chain in sickle hemoglobin and the other 16 pairs are found in the remaining three chains. The other 459 pairs of residues with hydrophobic interactions appear as intra-protein chains in HbS. Out of 459 pairs of intra-protein hydrophobic interactions, 252 pairs are associated with the β chain. The least hydrophobic interactions were observed in chain C(α) within 5 Å cutoff in sickled protein. LEU is found mostly (173 pairs) in contributing to the hydrophobic interactions in HbS. Out of 303 residue pairs involved in hydrophobic interactions in normal hemoglobin, 271 pairs are intra-protein hydrophobic interactions. Among them, LEU, VAL, PHE, ALA, TRP, PRO, TYR, and MET carry 89, 55, 40, 35,16, 15, 11, and 10 pairs of hydrophobic interactions, respectively, in the binding mechanism. The other 32 pairs of residues still carry the protein–protein hydrophobic interactions in normal hemoglobin. The β chain mostly interacted with hydrophobic potential in both proteins.

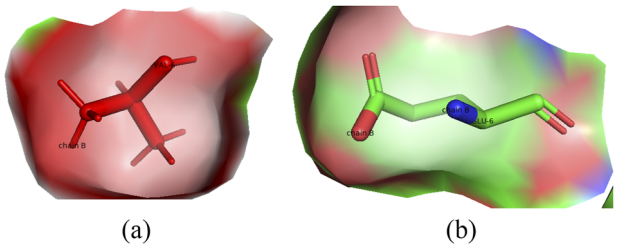


FIG. 7. (a) Hydrophobic surface of the sixth VAL of β chain of HbS. (b) Hydrophilic surface of sixth GLU of β chain of normal hemoglobin.

D. van der Waals (vdW) and electrostatic interactions

The vdW and electrostatic interactions are the non-bonded interactions that play important roles in the formation of a protein molecule. We have estimated such interactions in terms of non-bonded interaction energy for sickle and normal hemoglobin as presented in Tables I and II. We also carried out a quantitative estimation of the binding mechanism of the beta chain with other three chains (two α's and one β).

Comparing the data from Tables I and II, the sum of the vdW and electrostatic interaction energy of sickle hemoglobin (Table I) is higher than that of normal hemoglobin (Table II). We also investigated the sum of vdW and electrostatic interaction energy of the beta chain of both proteins and found a higher value for HbS than of normal hemoglobin at equilibrium. During N–C terminal SMD simulation of sickle hemoglobin with various spring constant k (500, 800, and 1100 kcal mol^{-1} nm^{-2}), the average of vdW and electrostatic interaction energy was observed very low around −600 kcal mol^{-1} (Table I). We also investigated the vdW and electrostatic interactions in the beta chain of normal hemoglobin by taking the values of spring constant k as 500, 800, and 1100 kcal mol^{-1} nm^{-2} in SMD simulation, and found much higher values around −1014.7904 to −4153.0754 kcal mol^{-1} (Table II). It indicates that the beta chain carries most of the vdW and electrostatic interaction energy in both proteins. It is also obvious that electrostatic potential energy in both HbS and A3N is much higher than vdW potential energy, which is similar to experimental results.

E. Salt bridges

Salt bridges are another important interaction in the formation of stable protein structures. Figure 8 represents the average of salt bridge occupancies in sickle and normal hemoglobin. We set the cut-off distance to 3.2 Å to analyze the salt bridges in different chains in the proteins. This study reveals that the residue pairs ASP94-LYS95, ASP94-LYS95, and GLU43-ARG92 have the occupancy percentage above 5 in the HbS (Fig. 8) (red) protein. The GLU43-ARG92

TABLE I. vdW and electrostatic interaction energy with suitable pulling velocity of 0.000 30 nm/ps in HbS protein.

For	vdW (kcal/mol)	Electro. (kcal/mol)	Sum (kcal/mol)
Equilibration HbS	−1663.0198	−7023.2588	−8686.2786
Spring constant(k) = 500 kcal mol^{-1} nm^{-2}	−133.1419	−556.4760	−689.6179
Spring constant(k) = 800 kcal mol^{-1} nm^{-2}	−113.7637	−597.4138	−711.1775
Spring constant(k) = 1100 kcal mol^{-1} nm^{-2}	−106.3517	−613.3389	−719.6906
Equilibration β chain	−440.8993	−1519.2683	−1960.1676

TABLE II. The average of the vdW and electrostatic interaction energy with pulling velocity of 0.000 30 nm/ps in A3N protein.

For	vdW (kcal/mol)	Electro.(kcal/mol)	Sum (kcal/mol)
Equilibration A3N	−1704.9993	−6505.7226	−8210.7221
Spring const.(k) = 500 kcal mol^{-1} nm^{-2}	−1014.7904	−4049.7745	−5064.5654
Spring const.(k) = 800 kcal mol^{-1} nm^{-2}	−1027.9575	−4235.3836	−5263.3418
Spring const.(k) = 1100 kcal mol^{-1} nm^{-2}	−1047.3884	−4153.0754	−5200.4645
Equilibration β chain	−425.2279	−1491.3647	−1916.5925

residue has the maximum contribution of salt bridge interaction in the HbS protein with occupancy of 5.80.

Furthermore, in normal hemoglobin, the average occupancy of salt bridges at a pulling velocity of 0.0003 nm/ps with spring constant in kcal mol^{-1} nm^{-2} is shown in Fig. 8 (blue). There are 32 pairs of salt bridges in binding the residues in this protein chain. Among them, eight pairs of residues (GLU121-LYS17, ASP94-LYS95, ASP85-LYS139, ASP94-LYS95, GLU101-ARG104, GLU101-ARG104, GLU30-LYS56, and ASP85-LYS139) have occupancy above 5. The maximum contribution of hydrophobic interaction was of GLU121-ARGLYS17 pairs in the A3N with occupancy of 6.30. ASP and GLU residues are found to be the most abundant in the salt bridge formation in all the chains. Furthermore, the salt bridge study estimated the presence of 30 salt bridges in sickle hemoglobin structure.

F. SMD of sickle and normal hemoglobin proteins

The SMD study of protein has been performed for both sickle and normal hemoglobin. Higher pulling velocities were avoided as they lead to loss of equilibration and significant error in the simulation. Likewise, very low pulling velocities are not effective in the SMD simulation. In the present work, the rupture force fluctuated greatly when the spring constant was larger than 1100 kcal mol^{-1} nm^{-2}. Meanwhile, the effect of lower spring

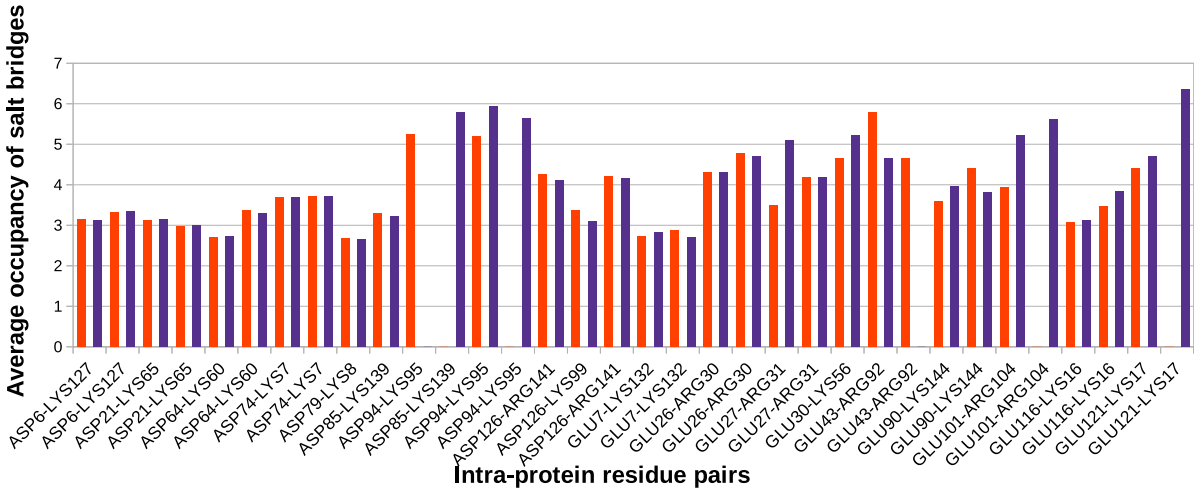


FIG. 8. Graph of average salt bridge frequency of residues in sickle and normal hemoglobin proteins.

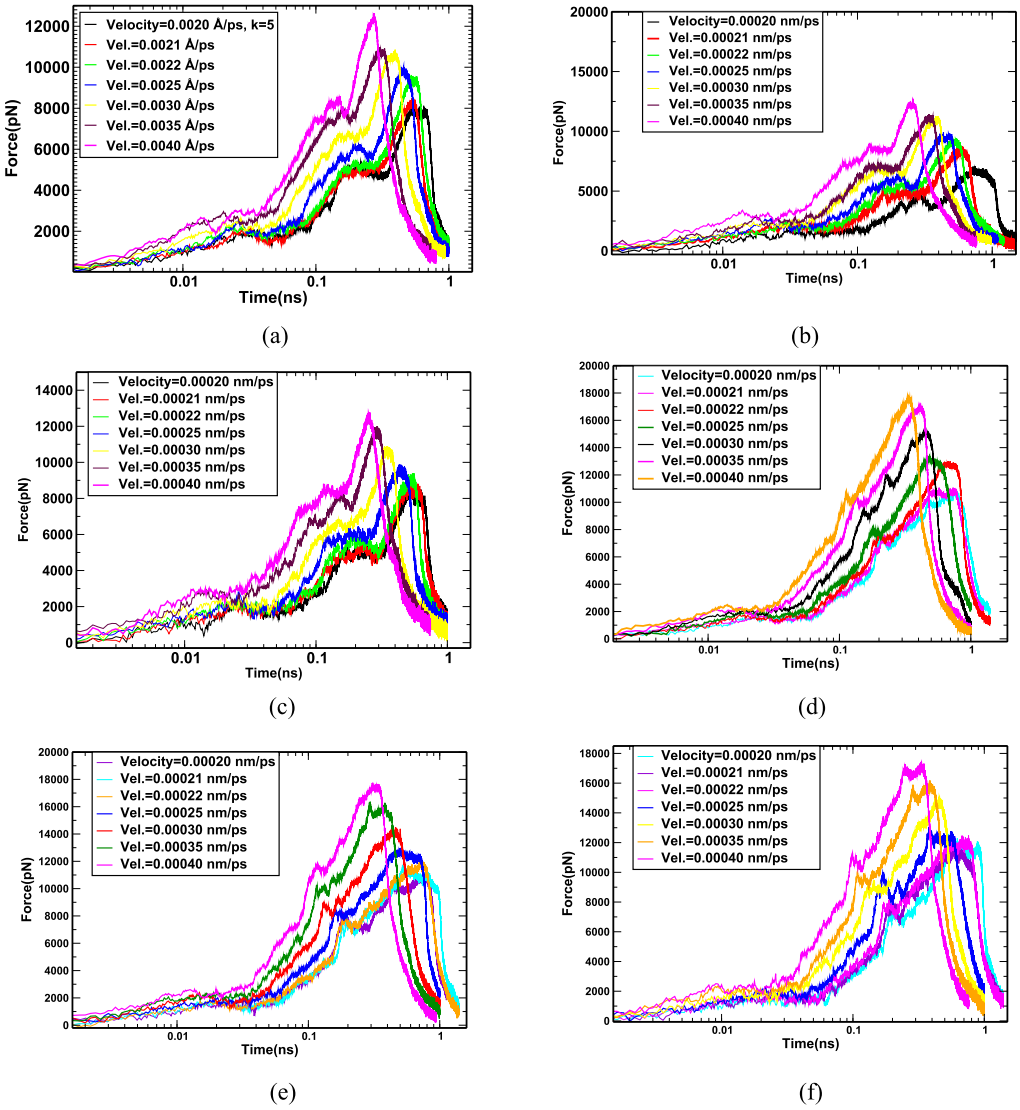


FIG. 9. SMD graph of (a) HbS at k = 500 kcal mol^{-1} nm^{-2}, (b) HbS at k = 800 kcal mol^{-1} nm^{-2}, (c) HbS at k = 1100 kcal mol^{-1} nm^{-2}, (d) A3N at k = 500 kcal mol^{-1} nm^{-2}, (e) A3N at k = 800 kcal mol^{-1} nm^{-2}, and (f) A3N at k = 1100 kcal mol^{-1} nm^{-2}.

TABLE III. Forces required for breaking the H-bond in sickle (HbS) and normal (A3N) hemoglobin with k = 500 kcal mol^{-1} nm^{-2}, k = 800 kcal mol^{-1} nm^{-2}, and k = 1100 kcal mol^{-1} nm^{-2}.

Force required (pN)	for HbS with k in kcal mol^{-1} nm^{-2}.			for A3N with k in kcal mol^{-1} nm^{-2}.		
Velocity (nm/ps)	For k = 500	For k = 800	For k = 1100	For k = 500	For k = 800	For k = 1100
0.000 20	8 120.70	7 073.74	8 908.22	10 809.90	12 005.00	10 425.40
0.000 21	8 470.74	8 759.14	8 984.77	10 881.50	12 577.10	12 820.80
0.000 22	9 621.09	9 471.70	9 397.36	12 421.30	11 761.80	11 336.90
0.000 25	10 045.30	9 902.31	9 881.79	13 396.80	12 951.60	13 032.60
0.000 30	10 858.20	11 286.50	10 900.00	15 276.50	14 473.10	15 140.90
0.000 35	10 972.00	11 420.10	12 015.00	17 113.30	16 359.40	16 187.20
0.000 40	12 643.40	12 646.80	12 770.50	17 676.70	17 753.70	17 434.10

constant value (less than 500 kcal mol^{-1} nm^{-2}) was not considered while confining the system. Therefore, in the present study, a spring constant between 500 and 1100 kcal mol^{-1} nm^{-2} was considered the best parameter in the SMD simulation for obtaining measurable and consistent rupture force. Figure 9 shows the graph of force required for breaking the hydrogen bond vs the time of simulation. The forces required in the constant velocity N–C terminal SMD simulation of normal and sickled hemoglobin have been estimated, and the results are presented in Figs. 9(a)–9(f).

We have executed the SMD simulation at different pulling velocities taking a fixed value of force constant. Furthermore, the force constant is also changed for the aforementioned velocities. In

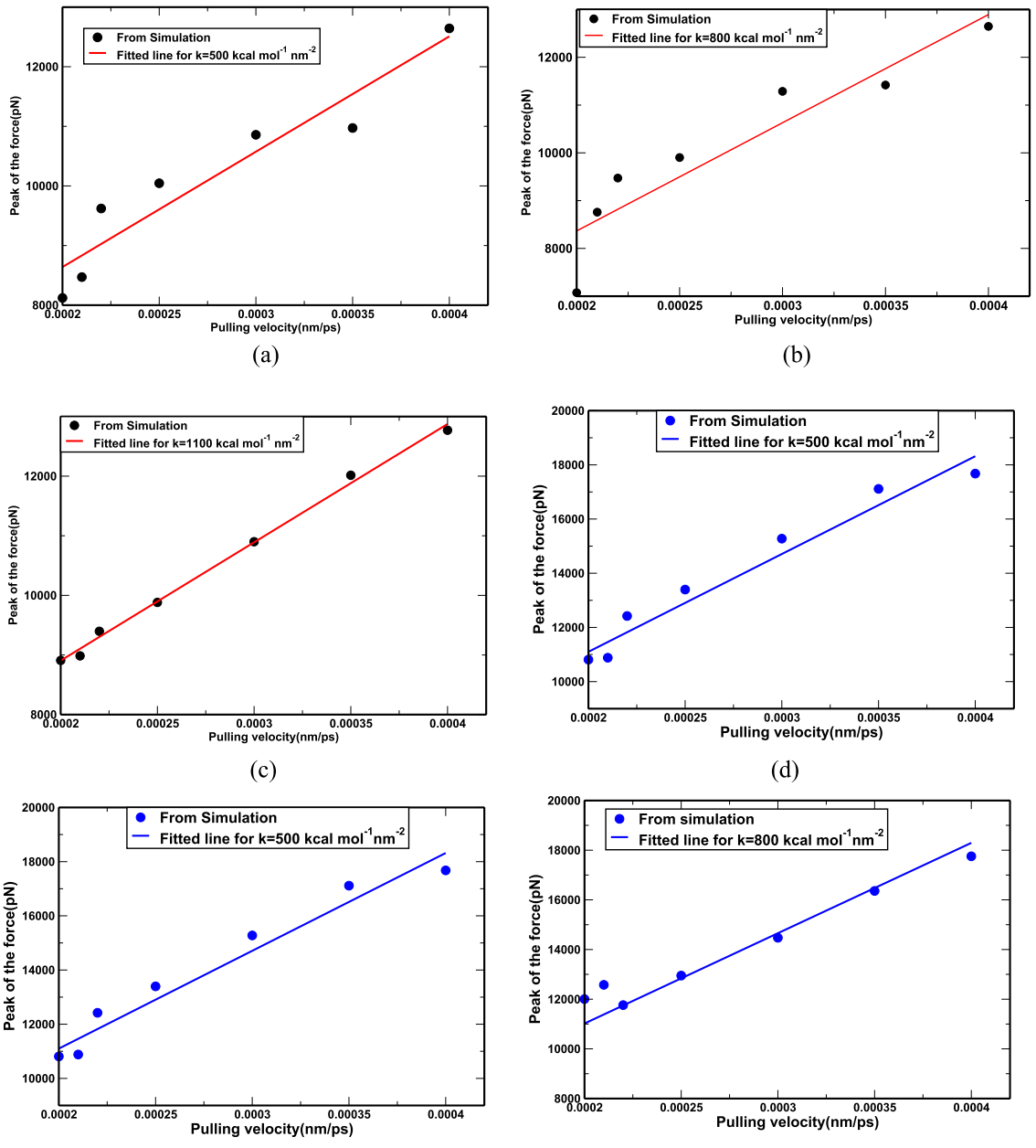


FIG. 10. Pulling velocity vs force required graph of (a) HbS with k = 500 kcal mol^{-1} nm^{-2}, (b) HbS with k = 800 kcal mol^{-1} nm^{-2}, (c) HbS with k = 1100 kcal mol^{-1} nm^{-2}, (d) A3N with k = 500 kcal mol^{-1} nm^{-2}, (e) A3N with k = 800 kcal mol^{-1} nm^{-2}, and (f) A3N with k = 1100 kcal mol^{-1} nm^{-2}.

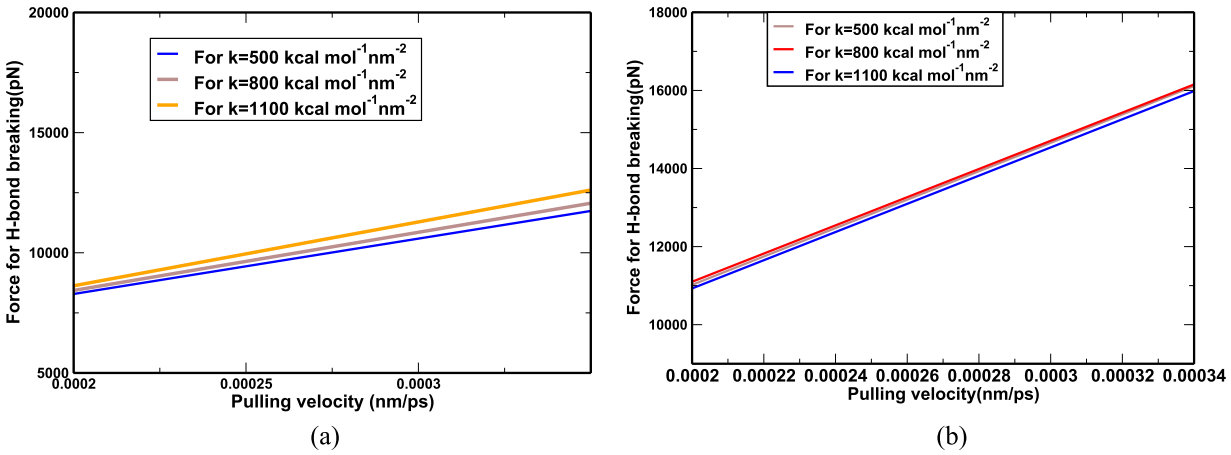


FIG. 11. Pulling velocity vs force required graph at different k of (a) HbS protein and (b) A3N protein.

Figs. 9(a) and 9(d), we observe that the peaks of the force required with much lower pulling velocity for both proteins are almost flat. At spring constant k = 500 kcal mol^{-1} nm^{-2}, the peak of the force is gradually increased as the pulling velocity is increased (Table III), i.e., the stiffness is higher at higher pulling velocity in both proteins. Similarly, the forces are increased as the pulling velocities are increased [Figs. 9(b), 9(c), 9(e), and 9(f)] for other two values of spring constant, i.e., k = 800 kcal mol^{-1} nm^{-2} and k = 1100 kcal mol^{-1} nm^{-2}, in both proteins (Table III).

The effects of pulling velocity on the rupture force for various spring constant k are shown in Figs. 10(a)–10(c). These graphs show that the rupture force has a linear correlation with the pulling velocities from 0.0002 to 0.0004 nm/ps in HbS. Therefore, the minimum pulling velocity was taken as 0.0002 nm/ps in the SMD simulation of HbS. Similarly, the effects of pulling velocity on the rupture force for various spring constant k (Table III) are shown in Figs. 10(d)–10(f), which also show the linear correlation with the pulling velocities from 0.0002 to 0.0004 nm/ps in normal hemoglobin. Therefore, minimum pulling velocity was taken as 0.0002 nm/ps in both cases of sickle and normal hemoglobin. It is found that as the spring constant k is increased the stiffness is increased in A3N.

We have plotted the graph between the pulling velocity and the average of the peak of the force needed for breaking the hydrogen bond with the values from Table III, which gives Figs. 11(a) and 11(b) for HbS and A3N, respectively.

With different values of k (500, 800, and 1100 kcal mol^{-1} nm^{-2}), the velocity vs force graph of HbS shown an increased slope for increasing velocity, i.e., the variation of force occurs with a uniform change in the velocity and the spring constant [Fig. 11(a)]. It also indicates that the higher the value of the spring constant, the higher is the force required, i.e., the stiffness is higher. We have also plotted the velocity vs force graph with different values of k (500, 800, and 1100 kcal mol^{-1} nm^{-2}) for A3N, and we obtained the same slope for each of the values of k. This indicates the variation of force with uniform change in velocity but not with spring constant k [Fig. 11(b)].

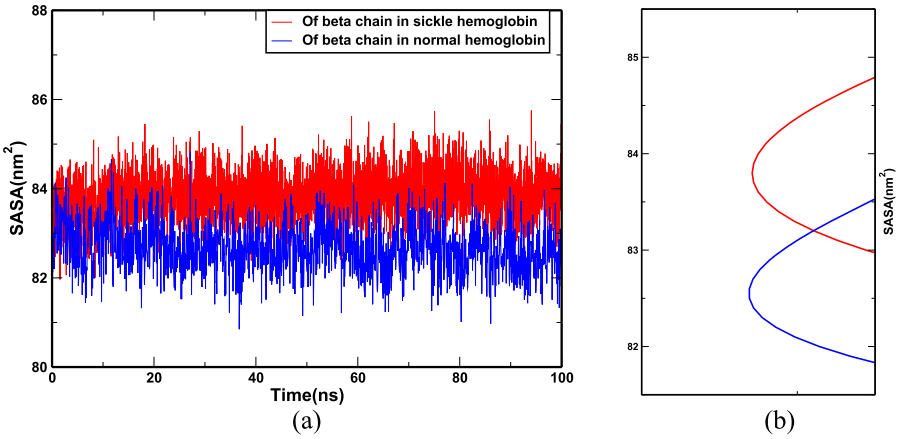


FIG. 12. (a) SASA values of the beta chain in sickle and normal hemoglobin. (b) The average SASA values of HbS (red) and A3N (blue).

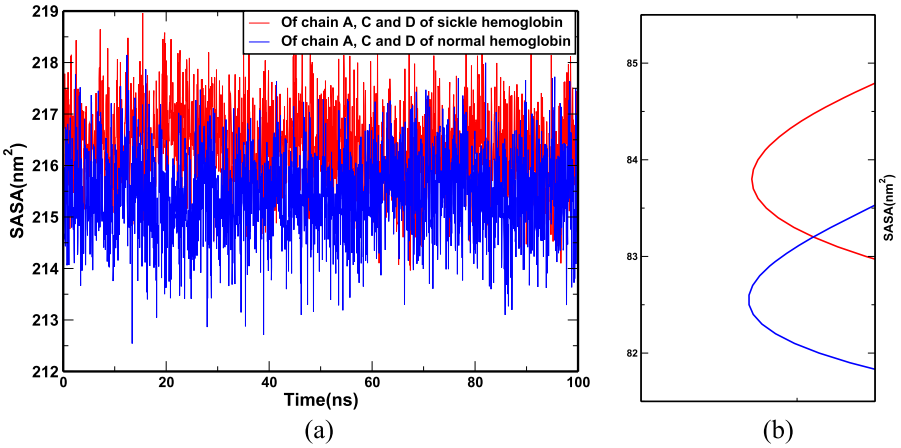


FIG. 13. (a) SASA values of A, C, and D chains in sickle and normal hemoglobin. (b) The average SASA values of A, C, and D chains of HbS (red) and A3N (blue).

G. Solvent accessible surface area (SASA)

Solvent accessible surface area (SASA) is the evaluation of the region of the molecular surface exposed enough to interact with solvent molecules. The contact surface area between one beta globin with other three (two alpha and one beta) is useful in examining the structural stability of the molecular complex. Here, we have estimated the SASA of one of the beta chains, ACD chains, and hemoglobin proteins of both sickled and normal hemoglobin. The SASA values of 100 ns SMD simulation have been compared to understand the binding affinity in sickle and normal hemoglobin proteins. An average value of solvent accessible surface area (SASA) of an individual molecule and their complex was determined by

$$\text{Contact area (S)} = (S_1(t) + S_2(t) - S_3(t))/2. \quad (5)$$

The SASA value increases with the time scale for different values of spring constant (k) and becomes constant, see Figs. 12(a), 13(a), 14(a), 15(a), and 15(b). At equilibrium, the SASA value of sickle hemoglobin protein (Table IV) is at 15.0520 nm^2; however, during the SMD simulation, the SASA value of the β chain of HbS decreased to around 0.9150 nm^{-2} [Fig. 12(a)]. In normal hemoglobin, at equilibrium, the SASA value is 0.2261 nm^{-2} (Table IV), and it decreased to around 0.8003 nm^{-2} during the SMD simulation of the β chain [Fig. 12(a)]. For higher values of k, the SASA value is also higher, which shows the hydrophilic behavior of the beta chain of sickle hemoglobin protein. There was a huge difference in SASA values of normal and sickle hemoglobin [Fig. 14(a)], which indicates the chances of a sharp reduction in interfacial tension (IFT) of normal hemoglobin than of sickle hemoglobin.[34] Figure 15(a) shows the SASA values of the beta chain in sickle hemoglobin with spring constants of k = 500 kcal mol^{-1} nm^{-2}, k = 800 kcal mol^{-1} nm^{-2}, and k = 1100 kcal mol^{-1} nm^{-2} as 0.8003, 0.7830, and 0.7908 nm^2, respectively. The estimated SASA values of the beta chain in normal hemoglobin for spring constant k = 500 kcal mol^{-1} nm^{-2}, k = 800 kcal mol^{-1} nm^{-2}, and k = 1100 kcal mol^{-1} nm^{-2} are 0.9120, 0.9150, and 0.9124 nm^2, respectively [Fig. 15(b)]. The average SASA of sickle hemoglobin decreased with the increase in k from 500 to 1100 kcal mol^{-1} nm^{-2} [Fig. 15(a)]. The SASA value of the beta chain of normal hemoglobin decreased with the increase in k from 800 to 1100 kcal mol^{-1} nm^{-2} [Fig. 15(b)]. However, the average SASA value observed at k = 500 kcal mol^{-1} nm^{-2} of normal hemoglobin is unexpected.

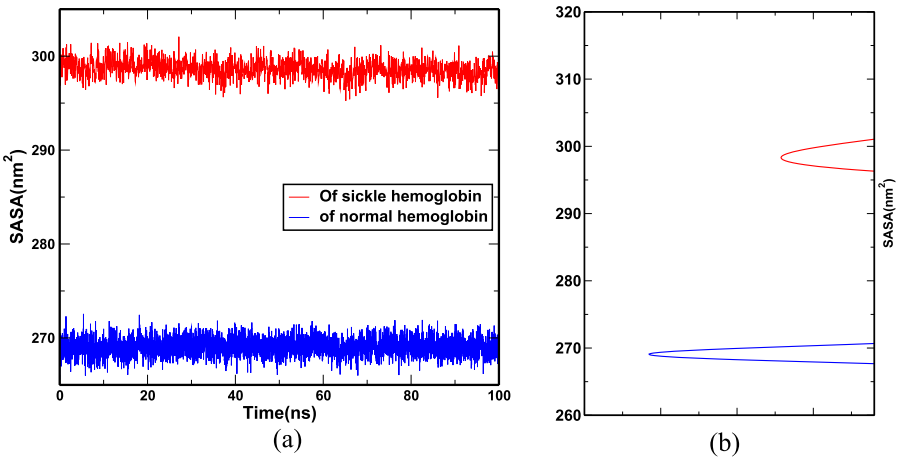


FIG. 14. (a) SASA values of sickle and normal hemoglobin. (b) The average SASA values of HbS (red) and A3N (blue).

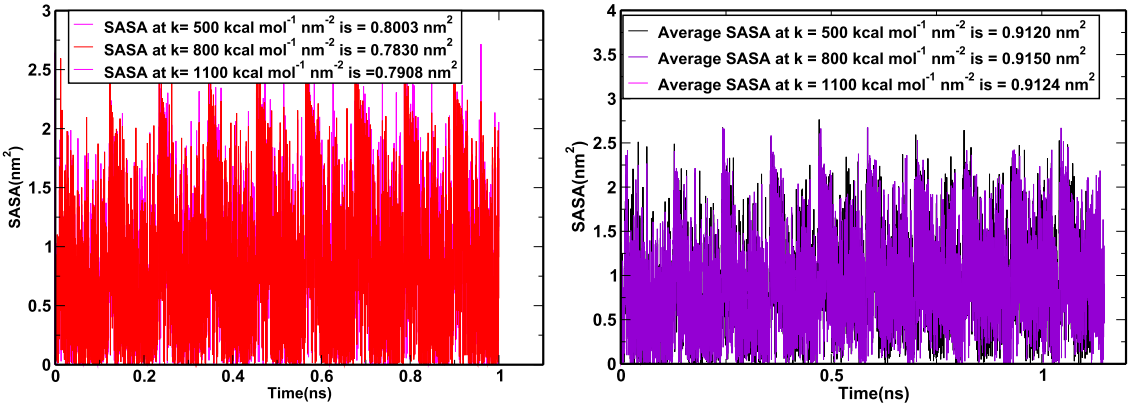


FIG. 15. SASA plot of equilibrated beta chain for different values of k in (a) sickle hemoglobin protein and (b) normal hemoglobin protein complex.

TABLE IV. Average SASA of β chain in HbS and A3N proteins with k = 500 kcal mol $^{-1}$ nm^{-2}.

	SASA (nm^2)			Net contact area (S) nm^2
Of	Chain B (S_1)	Chain ACD (S_2)	Protein (S_3)	= (S_1+S_2−s_3)/2
HbS	83.8514	215.3401	269.08.76	15.0520
A3N	82.6633	216.4071	298.6180	0.2261

IV. CONCLUSIONS

The structures of sickle and normal (deoxygenated human) hemoglobin achieved their stability at about 20 and 25 ns of time scale, respectively. Their corresponding relative RMSD values were stable at 0.036 and 0.035 nm. This study observed less flexibility of the atoms in the backbone of both sickle and normal hemoglobin as observed in a previous study by Prabhakaran and Johnson.[10] At the same time, the study indicates that the chain-B(β) is more strongly interacted by hydrogen bond and a hydrophobic interaction is involved in the hemoglobin protein. The hydrogen bond interaction in sickle hemoglobin is found more than in normal hemoglobin. In addition to them, the beta chain of sickle hemoglobin carries stronger vdW and electrostatic energy than the other three chains. The study also found that the hydrophobic interaction of normal hemoglobin is much less than that of sickle hemoglobin. The salt bridge analysis indicated a stronger interaction in normal hemoglobin than in sickle hemoglobin. Eight of the residue pairs containing mostly the glutamic acid and asperin residues strengthen this interaction in normal hemoglobin. However, in sickle hemoglobin, the strength in salt bridge interaction of these residues is reduced. It is seen that the glutamic acid and asperin amino acid mostly contribute to the salt bridge interaction in sickle hemoglobin.

The SMD simulation has confirmed that the forces for breaking the H-bond are higher for higher values of pulling velocity as well as spring constant used. As the stiffness depends on the value of k, we found higher stiffness for higher spring constant k as well as pulling velocity in both normal and sickle hemoglobin. The SASA estimation indicates that the sickle hemoglobin is more hydrophilic than normal hemoglobin as the SASA of sickle hemoglobin is much higher than that of normal hemoglobin. All these interactions play a certain role in transport property as well as interfacial tension of hemoglobin protein. Further study is anticipated for the estimation of interfacial tension and binding energy by umbrella sampling.

ACKNOWLEDGMENTS

J.P. acknowledges the University Grants Commission, Nepal, for Ph.D. Research Grant No. S & T-13/2076-2077. We acknowledge the TWAS for Research Grant No. RG 20-316. We thank Shyam P. Khanal and Rajendra P. Koirala for the fruitful discussion.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

J.P. carried out all the calculations, and N.P.A. conceived the idea for the study. All authors analyzed the results, read, and approved the final manuscript.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] J.-A. Ribeil, S. Hacein-Bey-Abina, E. Payen, A. Magnani, M. Semeraro, E. Magrin, L. Caccavelli, B. Neven, P. Bourget, W. El Nemer *et al.*, "Gene therapy in a patient with sickle cell disease," N. Engl. J. Med. **376**(9), 848–855 (2017).

[2] L. Pauling, H. A. Itano, S. J. Singer, and I. C. Wells, "Sickle cell anemia, a molecular disease," Science **110**(2865), 543–548 (1949).

[3] D. C. Rees, T. N. Williams, and M. T. Gladwin, "Sickle-cell disease," Lancet **376**(9757), 2018–2031 (2010).

[4] V. M. Ingram, "A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin," Nature **178**(4537), 792–794 (1956).

[5] J. T. Finch, M. F. Perutz, J. F. Bertles, and J. Döbler, "Structure of sickled erythrocytes and of sickle-cell hemoglobin fibers," Proc. Natl. Acad. Sci. U. S. A. **70**(3), 718–722 (1973).

[6] L. Luzzatto, "Sickle cell anaemia and malaria," Mediterr. J. Hematol. Infect. Dis. **4**(1), e2012065 (2012).

[7] O. S. Platt *et al.*, "Sickle cell anemia as an inflammatory disease," J. Clin. Invest. **106**(3), 337–338 (2000).

[8] S. Yesudasan, X. Wang, and R. D. Averett, "Molecular dynamics simulations indicate that deoxyhemoglobin, oxyhemoglobin, carboxyhemoglobin, and glycated hemoglobin under compression and shear exhibit an anisotropic mechanical behavior," J. Biomol. Struct. Dyn. **36**(6), 1417–1429 (2018).

[9] Y. Xue, S. Lofland, and X. Hu, "Thermal conductivity of protein-based materials: A review," Polymers **11**(3), 456 (2019).

[10] M. Prabhakaran and M. E. Johnson, "Molecular dynamics of sickle and normal hemoglobins," Biopolym.: Orig. Res. Biomol. **33**(5), 735–742 (1993).

[11] J. L. Maciaszek and G. Lykotrafitis, "Sickle cell trait human erythrocytes are significantly stiffer than normal," J. Biomech. **44**(4), 657–661 (2011).

[12] J. R. H. Tame and B. Vallone, "The structures of deoxy human haemoglobin and the mutant Hb Tyrα42His at 120 K," Acta Crystallogr., Sect. D: Biol. Crystallogr. **56**(7), 805–811 (2000).

[13] D. J. Harrington, K. Adachi, and W. E. Royer Jr., "The high resolution crystal structure of deoxyhemoglobin S," J. Mol. Biol. **272**(3), 398–407 (1997).

[14] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi *et al.*, "CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field," J. Chem. Theory Comput. **12**(1), 405–413 (2016).

[15] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," J. Chem. Phys. **79**(2), 926–935 (1983).

[16] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," J. Comput. Chem. **26**(16), 1781–1802 (2005).

[17] S. Jo, T. Kim, V. G. Iyer, and W. Im, "CHARMM-GUI: A web-based graphical user interface for CHARMM," J. Comput. Chem. **29**(11), 1859–1865 (2008).

[18] S. P. Khanal, Y. P. Kandel, and N. P. Adhikari, "Transport properties of zwitterion glycine, diglycine, and triglycine in water," AIP Adv. **9**(6), 065303 (2019).

[19] R. P. Koirala, R. Pokhrel, P. Baral, P. B. Tiwari, P. P. Chapagain, and N. P. Adhikari, "Structural insights into the repair mechanism of AGT for methyl-induced DNA damage," Biol. Chem. **402**(10), 1203–1211 (2021).

[20] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations," J. Comput. Chem. **4**(2), 187–217 (1983).

[21] A. Stradner, G. Foffi, N. Dorsaz, G. Thurston, and P. Schurtenberger, "New insight into cataract formation: Enhanced stability through mutual attraction," Phys. Rev. Lett. **99**(19), 198103 (2007).

[22] K. El Hage, F. Hédin, P. K. Gupta, M. Meuwly, and M. Karplus, "Valid molecular dynamics simulations of human hemoglobin require a surprisingly large box size," eLife **7**, e35560 (2018).

[23] R. P. Koirala, B. Thapa, S. P. Khanal, J. Powrel, R. P. Adhikari, and N. P. Adhikari, "Binding of SARS-CoV-2/SARS-CoV spike protein with human ACE2 receptor," J. Phys. Commun. **5**(3), 035010 (2021).

[24] G. Bussi, T. Zykova-Timan, and M. Parrinello, "Isothermal-isobaric molecular dynamics using stochastic velocity rescaling," J. Chem. Phys. **130**(7), 074101 (2009).

[25] D. González-Ruiz and H. Gohlke, "Targeting protein-protein interactions with small molecules: Challenges and perspectives for computational binding epitope detection and ligand finding," Curr. Med. Chem. **13**(22), 2607–2625 (2006).

[26] R. A. DiStasio, V. V. Gobre, and A. Tkatchenko, "Many-body van der Waals interactions in molecules and condensed matter," J. Phys.: Condens. Matter **26**(21), 213202 (2014).

[27] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, (Academic Press, 2002)

[28] J. E. Lennard-Jones, "Cohesion," Proc. Phys. Soc., London **43**(5), 461 (1931).

[29] B. L. Xiao, Y. N. Ning, N. N. Niu, D. Li, A. A. Moosavi-Movahedi, N. Sheibani, and J. Hong, "Steered molecular dynamic simulations of conformational lock of Cu, Zn-superoxide dismutase," Sci. Rep. **9**(1), 4353 (2019).

[30] F. Ercolessi, "A molecular dynamics primer," Spring College in Computational Physics, ICTP, Trieste, 1997, p. 19.

[31] W. D. Callister, *Materials Science and Engineering: An Introduction*, (John Wiley & Sons, Inc., UT, 2007)

[32] M. P. Allen and D. J. Tildesley, *Molecular Simulation of Liquids* (Clarendon, Oxford, UK, 1987)

[33] K. G. Tina, R. Bhadra, and N. Srinivasan, "PIC: Protein interactions calculator," Nucleic Acids Res. **35**(Suppl_2), W473–W476 (2007).

[34] M. P. Andersson, M. V. Bennetzen, A. Klamt, and S. L. S. Stipp, "First-principles prediction of liquid/liquid interfacial tension," J. Chem. Theory Comput. **10**(8), 3401–3408 (2014).

Binding mechanism of SARS-CoV-2 spike protein with human ACE2 receptor

Rajendra P Koirala
Tribhuvan University Kathmandu
Bidhya Thapa
Padhma Kanya Campus Tribhuvan University Kathmandu
Shyam P Khanal
Tribhuvan University Kathmandu
Jhulan Powrel
Tribhuvan University Kathmandu
Rajendra P Adhikari
Kathmandu University Dhulikhel
Narayan P Adhikari (narayan.adhikari@cdp.tu.edu.np)
Tribhuvan University Kathmandu https://orcid.org/0000-0003-4535-1368

Research Article

Keywords: SARS-COV 2, hACE2, Spike protein, Binding mechanism

Posted Date: September 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-71923/v1

Binding mechanism of SARS-CoV-2 spike protein with human ACE2 receptor

Rajendra P. Koirala[1], Bidhya Thapa[2], Shyam P. Khanal[1], Jhulan Powrel[1], Rajendra P. Adhikari[3] & Narayan P. Adhikari[1,*]

[*]**narayan.adhikari@cdp.tu.edu.np.**

[1]Central Department of Physics, Tribhuvan University, Kathmandu, Nepal
[2]Padma Kanya Multiple Campus, Tribhuvan University, Kathmandu, Nepal
[3]Department of Natural Sciences, Kathmandu University, Nepal

Abstract

SARS-CoV-2 virus interacts via C-terminal domain of spike protein to human cell receptor protein hACE2. Amino acid residues residing at the interface play vital role in binding of SARS-CoV-2 CTD to hACE2. The detailed atomic level investigation of interactions at binding interface of SARS-CoV-2 CTD/hACE2 provides indispensable information on better understanding of location for drug target. In the present work, we have studied the dynamical behaviour of the complex by analyzing the molecular dynamics (MD) trajectories. The major interacting residues of SARS-CoV-2 CTD and hACE2 have been identified by analyzing the non-bonded interactions such as hydrogen bondings, salt bridges, hydrophobic interactions, van der Waals interactions etc. Umbrella sampling method has been used to estimate the binding free energy for in-depth understanding of binding mechanism between virus protein and host receptor. The binding free energy difference, key residues at the interface, important atomic interactions and contact surface areas have been compared with the molecular complex of SARS-CoV and hACE2. Relatively larger contact surface area, more non-bonded interactions as well as greater binding free energy provide the evidence for favorable binding of SARS-CoV-2 with hACE2 receptor than SARS-CoV.

Introduction

Corona virus disease (COVID-19) pandemic, caused by severe acute respiratory syndrome (SARS)-like coronavirus (SARS-CoV-2), is a serious health concern for the global community [1, 2, 3]. Although the origin of the virus is still unclear, it has been spread all over the world threatening the human civilizations after its initial outbreak from China in December 2019. Till date, more than 23 millions infected population has been reported globally and more than eight hundred thousand people have lost their lives [4]. There is no approved drug or vaccine against the COVID-19, even though several antiviral drugs have been proposed and are also in clinical trials [5].

SARS-CoV-2 has more than 70 percent of structural similarity with SARS-CoV; most of the residues at binding interface are similar [7, 6]. SARS-CoV-2 similar with SARS-CoV and also other coronavirus utilizes human angiotensin converting enzyme 2 (hACE2) receptor to enter into human cell. This entry process is mediated by the spike(s) glycoprotein which are embedded in the capsid of SARS-CoV-2 [8]. The spike protein (S) of SARS-CoV-2 in most cases is cleaved by host into the subunits S1 and S2 which are responsible for receptor recognition and membrane fusion respectively [9]. Similar to SARS-CoV, C-Terminal Domain(CTD) of S1 subunit of spike protein in SARS-CoV-2

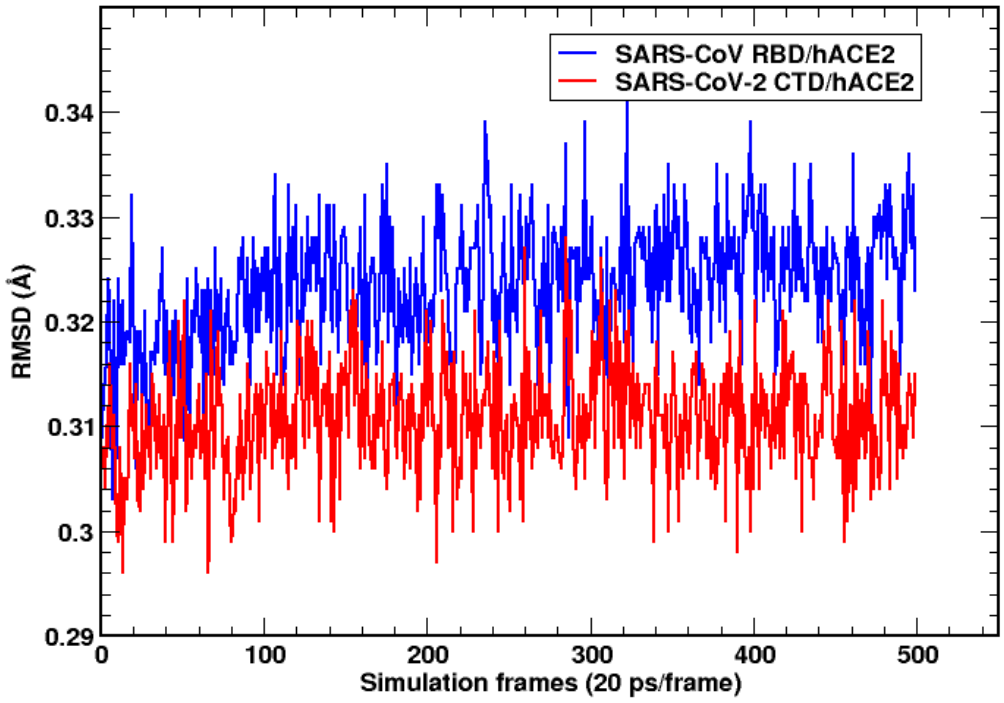
acts as receptor binding domain (RBD) [10, 11]. Even though both SARS-CoV and SARS-CoV-2 have same binding domain, the binding affinity of SARS-CoV-2 is greater than SARS-CoV [7, 12].
Immediately after the COVID-19 outbreak, several researches have been carried out to identify the nature and location of binding of SARS-CoV-2 CTD with hACE2 using static crystal structure [12, 13]. To our best knowledge molecular dynamics (MD) study of SARS-CoV-2 CTD/hACE2 has not been carried out to understand the binding mechanism during dynamics and also to estimate the free energy differences. The free energy calculation provides in-depth insight on the binding mechanism between the protein molecules [14]. There are several experimental techniques of measuring binding free energy such as isothermal titration calorimetry (ITC) [15], fluorescence resonance energy transfer (FRET) [16], nuclear magnetic resonance (NMR) [17], surface plasmon resonance (SPR) [18] and many others; however, these experimental methods are not easily accessible and time consuming. The computational approach can be the best complement for large scale investigations [19, 20, 21].
Out of many computational approach, umbrella sampling is one of the widely used method for the calculation of free energy in large molecular system [22, 23]. It improves the sampling system by designing and implementing the biasing potentials as a function of reaction coordinates [24, 25]. If an energy barrier exists in between two regions of configuration states, there may be poor sampling, despite the long simulation run being carried out. The applied biasing potential bridges such configuration states and makes it easier in searching local or global minima, which can be considered as the structurally favorable state in the molecular complex [26].
In this study, we have carried out molecular dynamics (MD) simulations for the comprehensive study of binding mechanism of SARS-CoV-2 CTD with hACE2. In addition, umbrella sampling method has been executed to estimate the binding free energy of SARS-CoV-2 CTD/hACE2. Required windows for the umbrella sampling have been taken from steered molecular dynamics (SMD) [27] simulations. In SMD, SARS-CoV-2 CTD has been translated, taking the hACE2 as the reference molecule. The quantitative estimation of binding affinity between the targeted molecules facilitates in silico-drug designing. Then, we have performed comparative study of various non-covalent atomic interactions at the binding interface of SARS-CoV-2 and SARS-CoV with hACE2. Furthermore, the binding free energy and contact surface area are compared of these complexes.

Results

MD simulations were performed to investigate the binding mechanisms of SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2 complexes by analyzing non-bonded interactions at the interface, contact surface area measurement and binding free energy calculation. This will be valuable to understand the entry mechanism of the virus to host receptor hACE2 and to provide molecular level insights on dynamic behavior of the complex.

Root Mean Square Deviation (RMSD). We checked the thermodynamic states of both systems (SARS-CoV-2 CTD/hACE2 complex and SARS-CoV RBD/hACE2 complex) to ensure whether the systems are conformationally stable or not. During 10 ns equilibration run, both systems were found conformationally stable after 2 ns simulation run. Then, RMSD of the backbones of both complexes were plotted with respect to time

dependent function of MD simulation as shown in Figure 1. From Figure 1, it is seen that the average RMSD of SARS-CoV-2 CTD/hACE2 complex was observed slightly smaller than that of SARS-CoV RBD/hACE2 complex, i.e., the complex of SARS-CoV-2 CTD and hACE2 is more stable than the complex of SARS-CoV RBD and hACE2.



[h]

Figure 1: RMSD of SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2.

Contact Surface Area. We also analyzed the contact surface area between the spike protein CTD/RBD and hACE2 receptor using MD trajectories. Contact area is the surface buried at the interface between two proteins which contributes to bind and stabilize the protein-protein complexes. Larger contact surface indicates greater stability of the structure [28]. Figure 2 shows the contact surface area for SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2 complexes. The contact surface area for SARS-CoV-2 CTD is more in comparison to SARS-CoV RBD indicating the greater binding affinity of SARS-CoV-2 with receptor. The estimated values of contact surface area are presented in Table 1. From the Table 1, it has been observed that SARS-CoV-2 CTD/hACE2 has larger contact area than SARS-CoV RBD/hACE2 by 77.0194 $\mathring{A}^2$.

Table 1: Calculation of contact surface area; A_1, A_2 & A_3 are the average solvent accessible surface area (SASA) of hACE2, SARS-CoV-2 CTD and their complexes; and $A_1^{'}$, $A_2^{'}$ & $A_3^{'}$ are the average solvent accessible surface area (SASA) of hACE2, SARS-CoV RBD and their complexes.

Complex	$SASA$ ($\mathring{A}^2$) for				
SARS-CoV-2 CTD/hACE2	hACE2 (A_1)	SARS-CoV-2 CTD (A_2)	complex(A_3)	$A_1 + A_2 - A_3$	Net contact area (A)[28]
	29100.3613	11227.9717	38549.0195	1779.3135	889.6568
SARS-CoV RBD/hACE2	hACE2 ($A_1^{'}$)	SARS-CoV RBD ($A_2^{'}$)	complex($A_3^{'}$)	$A_1^{'} + A_2^{'} - A_3^{'}$	Net contact area ($A^{'}$)
	29092.6152	10859.5420	38326.8825	1625.2747	812.6374

Non-bonded Interactions. Furthermore, we studied in details the hydrogen bonds,

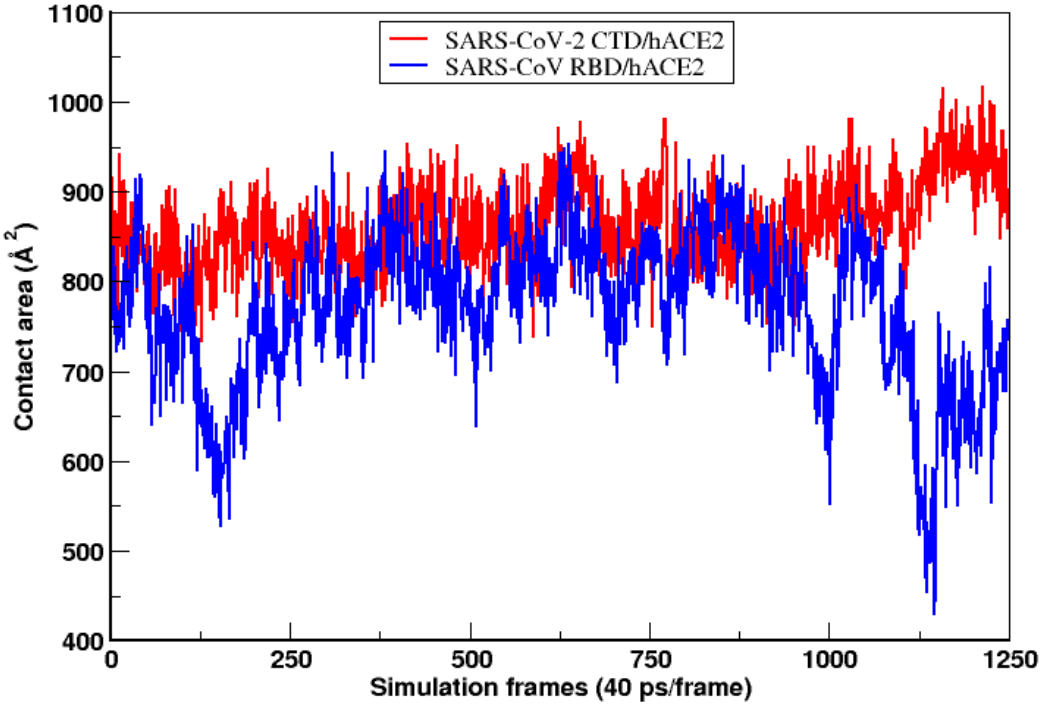


Figure 2: Comparison of contact surface area at the interface of SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2.

salt-bridges, hydrophobic, electrostatic and vdw interactions between the residues residing at the interface between SARS-CoV-2 CTD/hACE2; and also compared with SARS-CoV RBD/hACE2.

Hydrogen Bonds: Wang et al. (2020) and Lan et al. (2020) have studied the atomic interactions at the interface of static crystal structure of SARS-CoV-2 CTD/hACE2 complex [12, 13], whereas we have investigated the hydrogen bonds at the interface of two complexes by analyzing the MD simulations trajectories. The cut-off distance for hydrogen bond was taken 3.5 Å [12]. We monitored the time evolution of number of hydrogen bonds formed at the interface between SARS-CoV-2 CTD/hACE2 and also compared with that of SARS-CoV RBD/hACE2 as shown in Figure 3 (Also see supplementary Table S1). Hydrogen bonds were found to be consistently working through out the simulation. All potential interfacial hydrogen bonds formed in the SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2 during the 100 ns of simulations are as shown in Table 2. Total hydrogen bonds formed during the simulations were seen to be more in case of SARS-CoV; however the strength and life time of potential hydrogen bonds were found to be greater in case of SARS-CoV-2. Unlike SARS-CoV-2 CTD, some residues of SARS-CoV RBD form hydrogen bonds with the carbohydrates attached to the N-terminal domain of hACE2. The bond formed between VAL404 of SARS-CoV RBD with CARA-BMAN3 is significant.

There is no hydrogen bond formed between SER19 of hACE2 with any residues of SARS-CoV RBD in static structures [12, 13], however our results show two potential hydrogen bonds formed between main chain and side chain of SER19 of hACE2 with side chain of ASP463 of SARS-CoV RBD as in Figure 4(a). The hydrogen bonds formed at three key

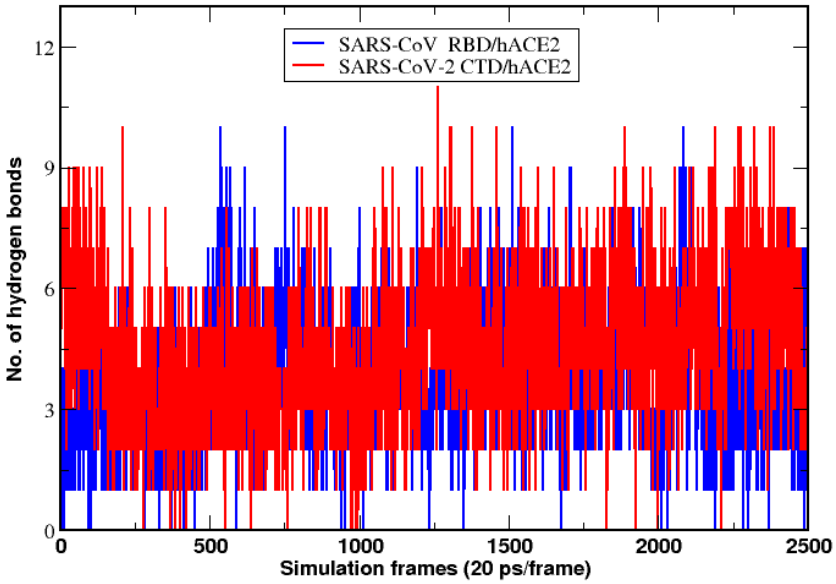


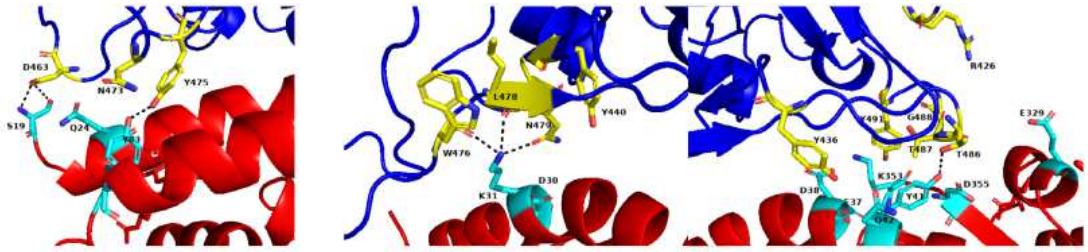
Figure 3: Comparison of time evolution of number of hydrogen bonds at the interfaces of SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2.

Table 2: Comparison of hydrogen bond occupancy percentage in SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2.

H-bond occupancy percentage	SARS-CoV-2 CTD/hACE2	SARS-CoV RBD/hACE2
< 1	47	80
1-10	19	22
10-20	1	2
20-30	2	4
30-40	4	1
40-50	1	1
50-60	2	1

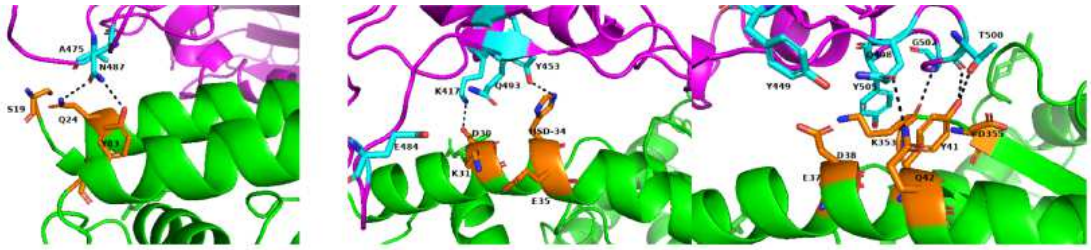
region of interface between SARS-CoV RBD/hACE2 at 0^{th} frame are shown in Figures 4(a), 4(b) and 4(c). We have observed some difference in the atomic interactions at the interface of both virus proteins and hACE2 than that of static crystal structure. The two different approaches might be the reason of variation in the number of interactions.

At the binding interface of SARS-CoV-2 CTD and hACE2 receptor, there are three key regions where most of the polar contacts between the interfacial residues take place. In region one, SER19, GLN24 and TYR83 of hACE2 form hydrogen bonds with ALA475 and ASN487 of SARS-CoV-2 CTD as in Figure 5(a) (also see supplementary Figures: S1(a), S2(a), S3(a) & S4(a)). In region two, there are interactions between the residues LYS417, TYR453 and GLN493 of SARS-CoV-2 CTD forming hydrogen bonds with ASP30, LYS31, HIS34 and GLU35 of hACE2 Figure 5(b) (also see supplementary Figures: S1(b), S2(b), S3(b) & S4(b)). Similarly, in region three, there is extensive network of hydrogen bonds between SARS-CoV-2 CTD residues GLY446, TYR449, GLY496, GLN498, THR500, ASN501, GLY502 and TYR505 with the hACE2 residues GLU37, ASP38, TYR41, GLN42, LYS353, ASP355 and ARG393 Figure 5(c) (also see supple-



(a) (b) (c)

Figure 4: Details of hydrogen bondings at the three key contact sites on binding interface between SARS-CoV RBD and hACE2 receptor. hACE2 receptor is shown in red and SARS-CoV RBD is shown in blue color. The contact residues which can form potential hydrogen bonds are shown as sticks & labelled by 1 residue letter (for 0^{th} frame).



(a) (b) (c)

Figure 5: Details of hydrogen bondings at the three key contact sites on binding interface between SARS-CoV-2 CTD and hACE2 receptor. hACE2 receptor is shown in green and SARS-CoV-2 CTD is shown in magenta color. The contact residues which can form potential hydrogen bonds are shown as sticks & labelled by 1 residue letter (for 0^{th} frame).

mentary Figures: S1(c), S2(c), S3(c) & S4(c)). Because of dynamical nature of our system, there is continuous breaking and formation of hydrogen bonds during the simulations. The variation of hydrogen bonds formed during the five representative frames of 100 ns simulation of SARS-CoV-2 has been shown in Figure 5 and supplementary Figures S1-S4.

Salt-bridges: In addition to extensive network of interfacial hydrogen bonds, another important contribution to protein-protein binding comes from salt-bridge interactions. MD trajectory analysis has shown four salt-bridges, having different bond length and strength, formed at the interface of SARS-CoV-2 CTD/hACE2. The salt-bridge formed between the residue LYS417 of SARS-CoV-2 CTD with ASP30 of hACE2 is found to be the strongest one among them owing to its short bond length. The remaining residues GLU484, LYS458 and ARG403 of SARS-CoV-2 CTD have formed salt-brigdes with LYS31, GLU23 and GLU37 of hACE2 respectively. In contrast, we found only one salt-bridge formed between ARG426 of SARS-CoV RBD and GLU329 of hACE2 which is much weaker than that of SARS-CoV-2 because of large bond length as in Figure 6. The mean contact score of each salt-bridge formed at both interfaces calculated using Pycontact (see supplementary Figure S5).

Hydrophobic Interactions: Hydrophobic patch created in the interface between SARS-CoV-2 CTD and hACE2 contributes to enhance the binding of protein complexes. In

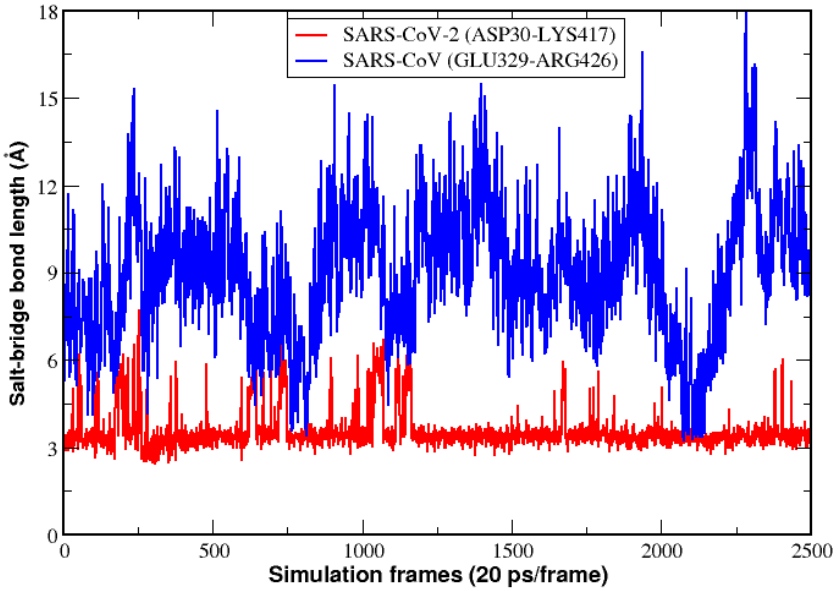


Figure 6: Comparison of time evolution of salt-bridge bond length at the interface of SARS-CoV RBD/hACE2 and SARS-CoV-2 CTD/hACE2.

SARS-CoV-2 CTD, PHE486 and TYR489 have hydrophobic contacts with the residues MET82, LEU79 and TYR89 of hACE2 (see supplementary Figure S6). It has been observed that PHE486 inserts into the hydrophobic groove created by the MET82, LEU79 and TYR89. This residue in SARS-CoV RBD is replaced by LEU472 being oriented in another direction has very weak interactions as compared to PHE486. Moreover, TYR489 has formed close and stable contact with PHE28 of hACE2. On the contrary, in SARS-CoV RBD only one residue TYR484 has formed strong and persistent hydrophobic contact with the TYR41 of hACE2 (see supplementary Figure S6).

Electrostatic and van der Waals (vdw) Interactions: The electrostatic and van der Waals (vdw) interactions in two complexes have been studied using NAMD Energy plugin package in VMD. Figure 7 depicts the comparative analysis of energy due to electrostatic and vdw interactions as a function of time for the two complexes. In the beginning of simulations, the electrostatic contributions of SARS-CoV-2 CTD/hACE2 was distinctly higher than SARS-CoV RBD/hACE2, however after 16 ns simulation time lapsed, the contributions were almost equal. In addition, the potential energy contributed by vdw interactions were consistently almost equal for both the systems throughout the simulations. It reveals that electrostatic and vdw interactions are almost equally contributed in binding both the complexes.

Free Energy. To investigate the energetic difference in binding of hACE2 with SARS-CoV-2 CTD and SARS-CoV RBD, the free energy differences have been estimated using Umbrella sampling technique. Umbrella windows were taken from the trajectories of SMD simulations. The interactions between the molecules in SARS-CoV-2 CTD/hACE2 were found terminated after traversing 9 Å distance away from the original position. To incorporate all interacting pathways, ten windows with 1 Å distance separation were taken for every successive window. On the other hand, the interactions between the

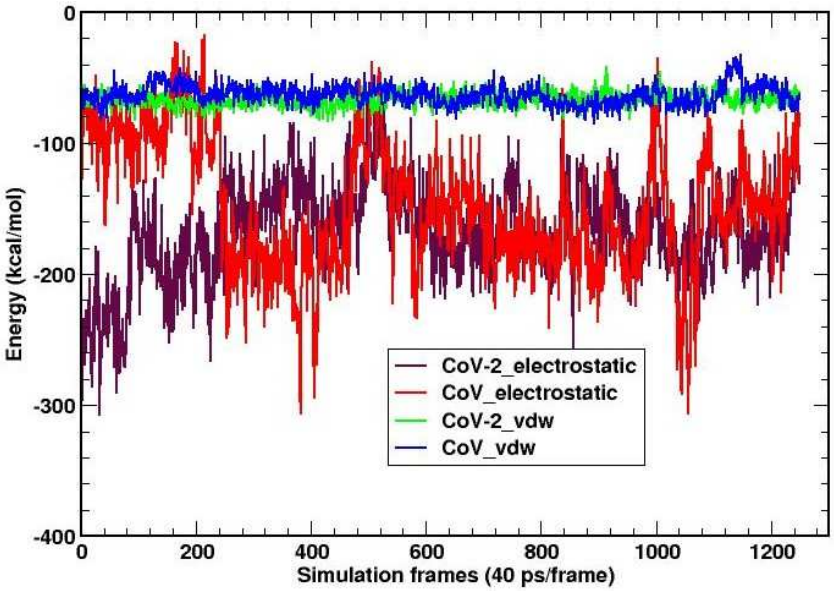


Figure 7: Comparison of electrostatic and vdw interactions of SARS-CoV-2 CTD and SARS-CoV RBD with hACE2.

molecules in SAR-CoV RBD/hACE2 were found ceased after traversing 5 Å distance from the original position. Therefore, six windows were prepared separating 1 Å distance away for every successive window. Figure 8 shows the change in free energy during the translation of virus CTD from active pocket of hACE2 for both complexes. The center of mass (COM) distance as a reaction coordinate allows us to track the free-energy changes for SARS-CoV-2 CTD in complex with hACE2 and SAR-CoV RBD in complex with hACE2 and compare the differences in the binding affinity for the two complexes. The SARS-CoV-2 CTD in complex with hACE2 is found to have the greater binding free energy of $\sim$ 1.91 kcal/mol compared to the SAR-CoV RBD in complex with hACE2. This, as well as the nature of the free-energy curve, provides an insight on binding mechanisms of the complexes.

Discussion

COVID-19 pandemic has seriously threatened public health throughout the globe. Since there is no approved drug till date to combat the people against the SARS-CoV-2 virus, more comprehensive study is essential through the various aspects at molecular level. The fundamental necessity is to understand the entry mechanism of the virus into the human cell, which is really helpful to discover the drug against the virus. To deal the entry mechanisms and dynamical characteristics of the virus cell in complex with hACE2 receptor, we used various computational techniques. C-Terminal Domain (CTD) of S1 subunit of spike protein, being the active interacting region, has been taken into consideration in SARS-CoV-2. We performed the comparative analysis of the key residues and atomic interactions responsible for the binding of the SARS-CoV-2 CTD and SARS-CoV RBD with human ACE2 receptor.

Estimation of structural variation during the simulation is the foremost judgement of molecular stability in molecular dynamics study. RMSD is the measure of stability of molecular structure in the cellular environment. Well equilibrated system with consistent

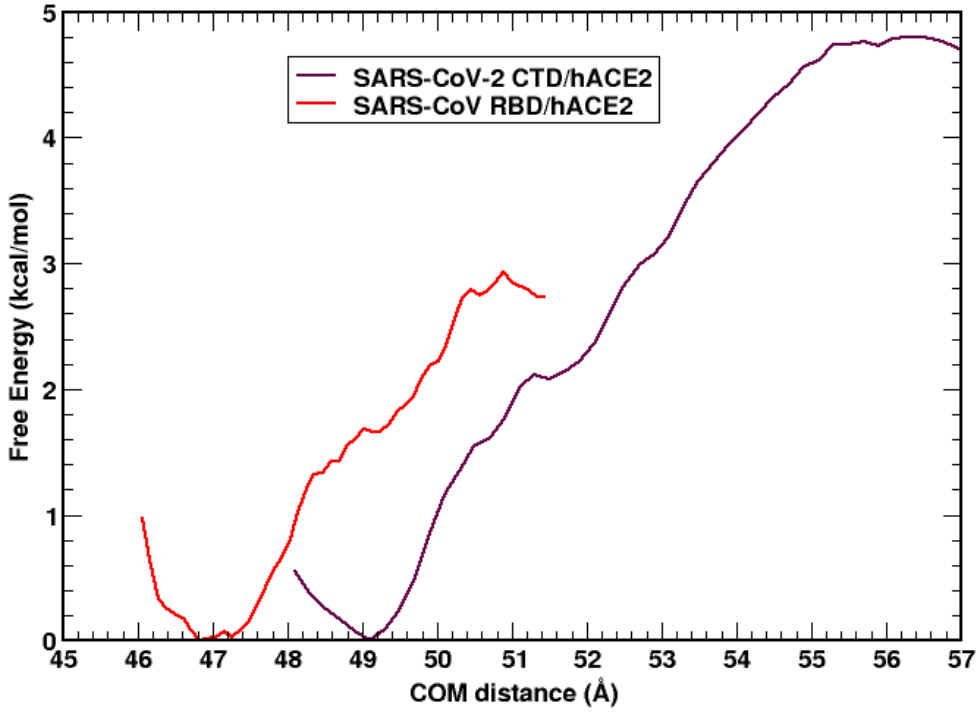


Figure 8: Free energy curve for SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2 during the translation of spikes CTD/RBD from hACE2.

RMSD ensures us to proceed for the further study of binding affinity and energy variations of the molecular complexes [29]. Moreover, contact surface area between the molecules identifies the binding strength of the complex. Therefore, we have obtained the contact surface area of both complexes calculating the SASA. SASA has been determined from time evolution data generated from the 100 ns NVT run. Then, average value of contact area for both the systems have been presented in Table 1 and are interpreted graphically in Figure 2. Larger contact surface area in SARS-CoV-2 CTD/hACE2 complex depicts the stronger binding of this complex than that of SARS-CoV RBD/hACE2 [30].
Our results show considerable similarity in the binding sites, interfacial residues and important atomic interactions in both viral protein receptor binding domain (i.e., SARS-CoV-2 CTD and SARS-CoV RBD). However, there is some structural variation in a loop between two structures in the binding region and some residues at the binding sites are different. This facilitates more and stronger atomic contacts between SARS-CoV-2 CTD and hACE2 interface and thereby enhancing its binding affinity. Polar residues residing at the interface form an extensive network of hydrogen bonds and salt-bridge interactions [31, 32, 33, 34]. Our study reveals that interfacial hydrogen bonds, salt-bridges and hydrophobic interactions play an important role in the binding of SARS-CoV-2 CTD to host cell receptor. Furthermore, comparative analyses of the binding mechanism of two

viral proteins with hACE2 show that binding affinity of SARS-CoV-2 is greater than that of SARS-CoV. Notably, more residues are engaged in the binding of SARS-CoV-2 CTD with hACE2. We find the greater number of potential hydrogen bonds formed in the case of SARS-CoV-2 CTD which contributes to higher binding affinity. More and stronger salt-bridges formed in case of SARS-CoV-2/hACE2 establish stronger binding to the receptor. Additionally, we observe hydrophobic interactions are stronger in case of SARS-CoV-2 which also contribute to enhanced binding.
The contributions of electrostatic and vdw contacts are significant to form a stable protein-protein complex [35, 36]. The potential energy in binding the virus CTD/RBD and host receptor are compared in both the systems. Though, initially the electrostatic energy is observed relatively larger in SARS-CoV-2 CTD/hACE2 than that of SARS-CoV RBD/hACE2, the dynamical results show almost equal contributions in both the complexes. This shows that the contributions of hydrogen bonds, salt bridges and hydrophobic interactions are responsible to provide the greater binding strength in SARS-CoV-2 CTD/hACE2.
The binding mechanisms of the complexes are further analyzed estimating free energy differences from umbrella sampling method. SMD trajectories are taken for the appropriate samples that ensure the sufficient overlapping on windows [37]. In SMD, the virus CTD/RBD are pulled upto that distance, beyond which no interactions persists. We find the interactions of molecules in complex SARS-CoV RBD/hACE2 have been terminated after the displacement of RBD by 5 Å from host receptor, whereas the interactions sustain upto 9 Å displacement from the initial position in SARS-CoV-2 CTD/hACE2. Comparisons of free energy of two complexes has provided the insight of bonding affinity between the virus CTD/RBD and hACE2 molecules. The greater free energy difference between SARS-CoV-2 CTD in complex with hACE2 depicts the stronger binding strength than the complex of SARS-CoV RBD and hACE2. As the further investigation, we plan to calculate the solvation free energy of SARS-CoV-2 and SARS-CoV molecule in the aqueous environment.

Methods

System Set Up. Two molecular structures, PDB IDs 6LZG and 2AJF, were taken for the molecular dynamics simulations. The PDB ID 6LZG contains the complex of SARS-CoV-2 CTD and hACE2 receptor protein (i.e., SARS-CoV-2 CTD/hACE2 complex) and that of PDB ID 2AJF contains the complex of SARS-CoV RBD and hACE2 receptor protein (i.e., SARS-CoV RBD/hACE2 complex) [38]. CHARMM-GUI [39] was used to create the pdb and psf files of these complexes. Then, both the complex structures were solvated using TIP3P [40] water and electrically neutralized by adding NaCl. A cubical box of dimensions $144 \times 144 \times 144$ Å^3 was prepared for NVT simulation of the complex SARS-CoV-2 CTD/hACE2 and another cubical box of dimensions $131 \times 131 \times 131$ Å^3 was prepared for NVT simulation of the complex SARS-CoV RBD/hACE2. Further, two equal sized orthorhombic simulation boxes were prepared in order to estimate the free energy differences of above complexes by changing the dimensions to $250 \times 90 \times 90$ Å^3 and electrically neutralized by adding NaCl.

Molecular Dynamics Simulation. All atom molecular dynamics (MD) simulations were performed using NAMD [27] simulation package. The CHARMM36m [41] force

field was used for each simulations. The Particle Mesh Ewald (PME) [42] was used for the long-range interactions with a 12.0 $\mathring{A}$ non-bonded cut-off. The energy minimization was performed for 10,000 steps, using the conjugate gradient and line search algorithm [43, 44]. The system was then equilibrated at 310 K for 10 ns with harmonically restrained heavy atoms taking 1 fs time step. Finally, the production run was propagated for 100 ns simulation under NVT conditions by using Langevin dynamics with a damping constant of 1 ps^{-1} taking time step of 2 fs.

Steered Molecular Dynamics and Umbrella Sampling. To perform the umbrella sampling, sample windows were chosen from steered molecular dynamics (SMD) trajectories. During SMD, CTD/RBD of SARS-CoV-2 CTD/SARS-CoV RBD were pulled correspondingly towards the negative x-direction with constant velocity pulling method of velocity $0.00005\ \mathring{A}/fs$. In this process, the alpha carbons of hACE2 protein were taken as the fixed atoms and alpha carbons in CTD/RBD part of spike protein of the systems were taken as the dummy atoms; and were pulled from their center of mass (COM) along the negative x-direction with constant velocity ($\vec{v} = d\vec{x}/dt$) in water and ions environment so that the SMD atom experiences the force $\vec{F(t)} = k(\vec{v}\ t - \vec{\Delta x})$, providing the external potential energy,

$$U(x,t) = \frac{1}{2}k(\vec{v}\ t - \Delta\vec{x}.\hat{n})^2 \tag{1}$$

where, k ($= 5\ kcal\ mol^{-1}\mathring{A}^{-2}$) is the spring constant and gives the stiffness of the applied harmonic restraining force, and $\Delta\vec{x}\ (t) = \vec{x}\ (t) - \vec{x}_0$, is the displacement of SMD molecules from initial position $\vec{x}_0$ to instantaneous position $\vec{x}\ (t)$ and $\hat{n}$ is the unit vector along the direction of pulling.
Umbrella sampling was performed to investigate the free energy difference during the translation of SARS-CoV-2 CTD from hACE2 protein for system SARS-CoV-2 CTD/hACE2; and identical condition is applied for system SARS-CoV RBD/hACE2. SMD trajectories were used to select the appropriate windows. Identifying the information on the termination of molecular interactions from SMD, we estimated the number of umbrella windows in both the systems. Ten windows were prepared in SARS-CoV-2 CTD/hACE2 and six windows were prepared for SARS-CoV RBD/hACE2 complexes. Every successive window was taken from the SMD trajectories during the translation of 1 Å along the negative x-direction. The window size ensures the sufficient overlapping of successive windows to cover the entire reaction coordinate space. The reaction coordinate was chosen as the distance between the center of mass (COM) of hACE2 and CTD/RBD spike along the negative x-axis. To make the necessary overlapping reaction coordinates, a bias potential V(x) was used to force the system to fluctuate in coordinate space, which is given by,

$$V(x) = \frac{1}{2}k_i(x - x_0)^2 \tag{2}$$

where x_0 is the harmonic constraint defining a center of window i (i = 1 to 10 for for SARS-CoV-2 and 1 to 6 for SARS-CoV), and force constant k_i is the window width. We used the force constant of 1.5 $kcal\ mol^{-1}\mathring{A}^{-2}$.

Data Analysis. Visual Molecular Dynamics (VMD) [45] and Pymol [46] were used to visualize as well as generate images of the complex structures . VMD analysis tools were

used to identify and analyze non-bonded interactions by using the simulation trajectories. The NAMD energy plugin, available in VMD, was used to calculate the non-bonded interaction energy contributions. Pycontact [47] software package was used to analyzed the hydrophobic interactions and salt bridges between the targeted protein residues in CTD/RBD of spike protein and ACE2. Weighted Histogram Analysis Method (WHAM) program [48] was used to estimate the free energy from umbrella sampling simulation. The free energy calculation of large molecular system is generally computationally demanding. This method minimizes the statistical errors as well as increases the efficiency of computational simulation. Moreover, it has the advantage of multiple overlapping of probability distributions for obtaining better estimation of free energy calculations.

References

[1] Yan, R. et al. Structural basis for the recognition of the 2019-nCoV by human ACE2. *Science* **367**, 1444–1448 (2020).

[2] Guan, W. et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).

[3] Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* (2020).

[4] https://covid19.who.int/table.

[5] Folegatti, P. M. et al. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. *The Lancet* (2020).

[6] Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* (2020).

[7] Shang, J. et al. Clinical characteristics of coronavirus disease 2019 in China. *Nature* **581**, 221–224 (2020).

[8] Lu, G., Wang, Q. & Gao, G. F. Bat-to-human: spike features determining host jumpof coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol.* **23**, 8, 468–478 (2015).

[9] Lai M. Coronaviridae. *Fields virology* 1305–1318 (2007).

[10] Lu, G. et al. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. *Nature* **500**, 7461, 227–231 (2013).

[11] Li, F. et al. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **309**, 5742, 1864–1868 (2005).

[12] Wang, Q. et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* (2020).

[13] Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 7807, 215–220 (2020).

[14] Peng, Y. et al. Predicting protein–DNA binding free energy change upon missense mutations using modified MM/PBSA approach: SAMPDI webserver. *Bioinformatics* **34**, 5, 779–786 (2018).

[15] Velázquez-Campoy, A et al. Isothermal titration calorimetry. *Curr. Protoc. Cell Biol.* **23**, 1, 17–8 (2004).

[16] Hillisch, A., Lorenz, M. & Diekmann, S. Recent advances in FRET: distance determination in protein–DNA complexes. *Curr. Opin. Struct. Biol.* **11**, 2, 201–207 (2001).

[17] Campagne, S., and Gervais, V. & Milon, A. Nuclear magnetic resonance analysis of protein–DNA interactions. *J. R. Soc. Interface.* **8**, 61, 1065–1078 (2011).

[18] Teh, H., et al. Characterization of protein- DNA interactions using surface plasmon resonance spectroscopy with various assay schemes. *Biochemistry* **46**, 8, 2127–2135 (2007).

[19] Donald, J. E., Chen, W. W. & Shakhnovich, E. I. Energetics of protein–DNA interactions. *Nucleic Acids Res.* **35**, 4, 1039–1074 (2007).

[20] Jones, S. et, al. Protein-DNA interactions: a structural analysis. *J. Mol. Biol.* **287**, 5, 877–896 (1999).

[21] Zhang, C. et, al. A knowledge-based energy function for protein- ligand, protein-protein, and protein- DNA complexes. *J. Med. Chem.* **48**, 7, 2325–2335 (2005).

[22] Torrie, G. M. & Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **23**, 2, 187–199 (1977).

[23] Kästner, J. Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 6, 932–942 (2011).

[24] Zheng, L., Chen, M. & Yan W. Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. *Proc. Natl. Acad. Sci.* **105**, 51, 20227–20232 (2008).

[25] Yang, M., et al. Combine umbrella sampling with integrated tempering method for efficient and accurate calculation of free energy changes of complex energy surface. *J. Chem. Phys.* **141**, 4, 07B618_1 (2014).

[26] Yang, Y., et al. Efficient sampling over rough energy landscapes with high barriers: A combination of metadynamics with integrated tempering sampling. *J. Chem. Phys.* **144**, 9, 094105 (2016).

[27] Phillips, J. et, al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 16, 1781–1802 (2005).

[28] Zou, X. et al. Recognition of methylated DNA through methyl-CpG binding domain proteins. *Nucleic Acids Res.* **40**, 6, 2747–2758 (2012).

[29] Kufareva, I. & Ruben, A. Methods of protein structure comparison. Homology Modeling , 231–257 (2011).

[30] Ma, B. et al. Protein–protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc. Natl. Acad. Sci.* **100**, 10 5772–5777 (2003).

[31] Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discov.* **2**, 7 527–541 (2003).

[32] Chen, J., Sawyer, N. & Rayan, L. Protein–protein interactions: General trends in the relationship between binding affinity and interfacial buried surface area. *Protein Science* **22**, 4 510–515 (2013).

[33] Jones, S. & Thornton, J. M. Analysis of protein-protein interaction sites using surface patches. *J. Mol. Biol.* **272**, 1 121–132 (1997).

[34] Xu, D., Tsai., C. & Nussinov, R. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Eng.* **10**, 9 999–1012 (1997).

[35] Jones, S. et al. Using electrostatic potentials to predict DNA-binding sites on DNA-binding proteins. *Nucleic Acids Res.* **31**, 24, 7189–7198 (2003).

[36] DiStasio Jr, R. A., Gobre, V. V. & Tkatchenko, A. Many-body van der Waals interactions in molecules and condensed matter. *J. Phys. Condens. Matter* **26**, 21 213202 (2014).

[37] Isralewitz, B. et al. Steered molecular dynamics investigations of protein function. *J. Mol. Graph. Model.* **19**, 1, 13–25 (2001).

[38] Berman, H. et, al. The protein data bank. *Nucleic Acids Res.* **28**, 1, 235–242 (2000).

[39] Lee, J. et, al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 1, 405–413 (2016).

[40] Jorgensen, W. et, al. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 2, 926–935 (1983).

[41] Huang, J. et, al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **14**, 1, 71–71 (2017).

[42] Harvey, M. J., Giupponi, G. & Fabritiis, G. De. ACEMD: accelerating biomolecular dynamics in the microsecond time scale. *Nature methods* **5**, 6, 1632–1639 (2009).

[43] Khanal, S. P., Kandel, Y. P. & Adhikari, N. P. Transport properties of zwitterion glycine, diglycine, and triglycine in water. *AIP Advances* **9**, 6, 065303–10 (2019).

[44] Koirala, R. P. , Bhusal, H. P. , Khanal, S. P. & Adhikari, N. P. Effect of temperature on transport properties of cysteine in water. *AIP Advances* **10**, 2, 025122 (2020).

[45] Humphrey, W. et al. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 1, 33–38 (1996).

[46] Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).

[47] Scheurer, M. et al. PyContact: Rapid, customizable, and visual analysis of noncovalent interactions in MD simulations. *Biophysical journal* **114**, 3, 577–583 (2018).

[48] Kumar, S. et al. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 8, 1011–1021 (1992).

Acknowledgements: RPK & SPK acknowledge the partial financial support from Nepal Academy of Science and Technology (NAST). NPA acknowledges the UGC Award no. CRG-73/74-S&T-01 and TWAS research grants RG 20-316. We acknowledge the computing facilities of Supercomputer Centre Kathmandu University, which was established with equipment donated by CERN and the Arkansas High Performance Computing Center which is funded through multiple National Science Foundation grants and the Arkansas Economic Development Commission.

Competing Interests: The authors declare that they have no competing financial interests.

Figures

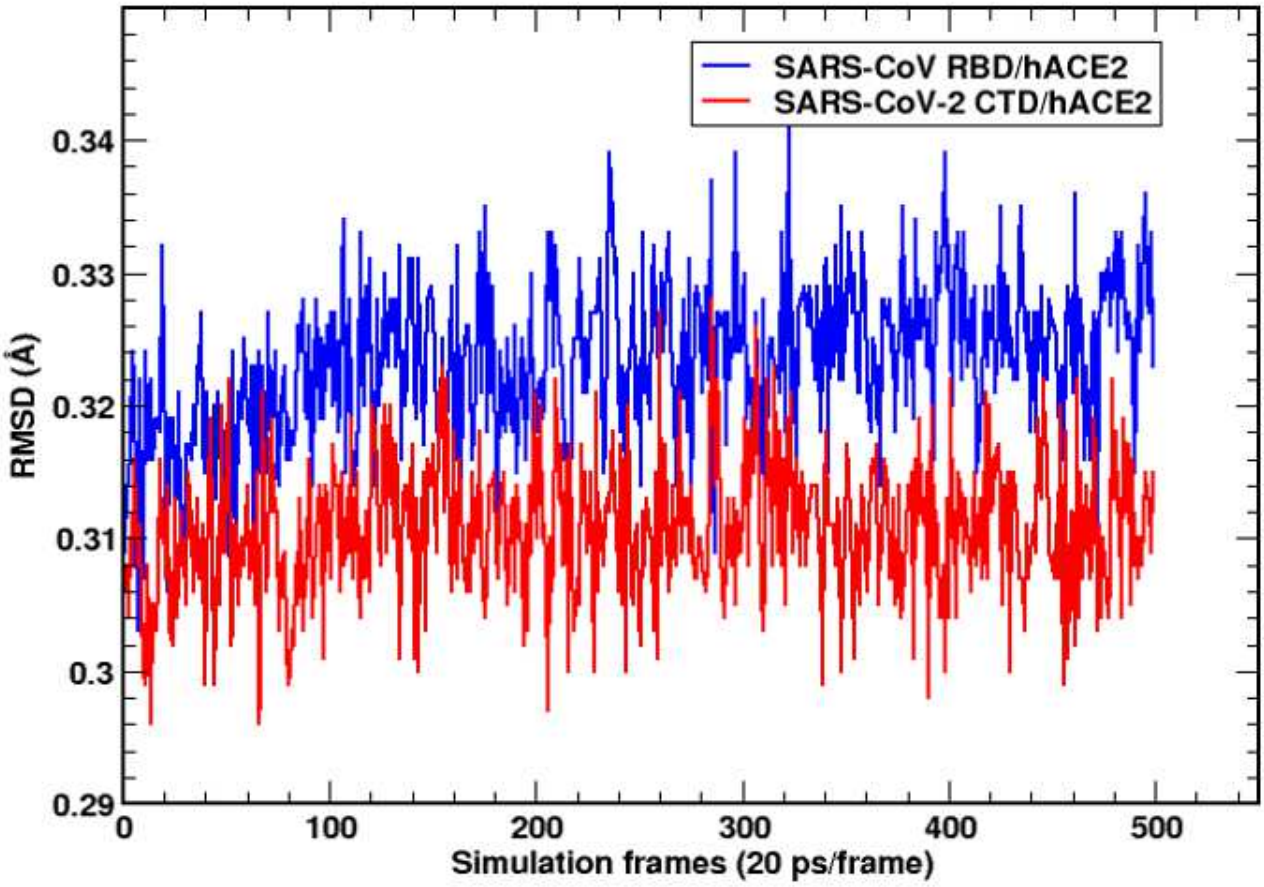


Figure 1

RMSD of SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2.

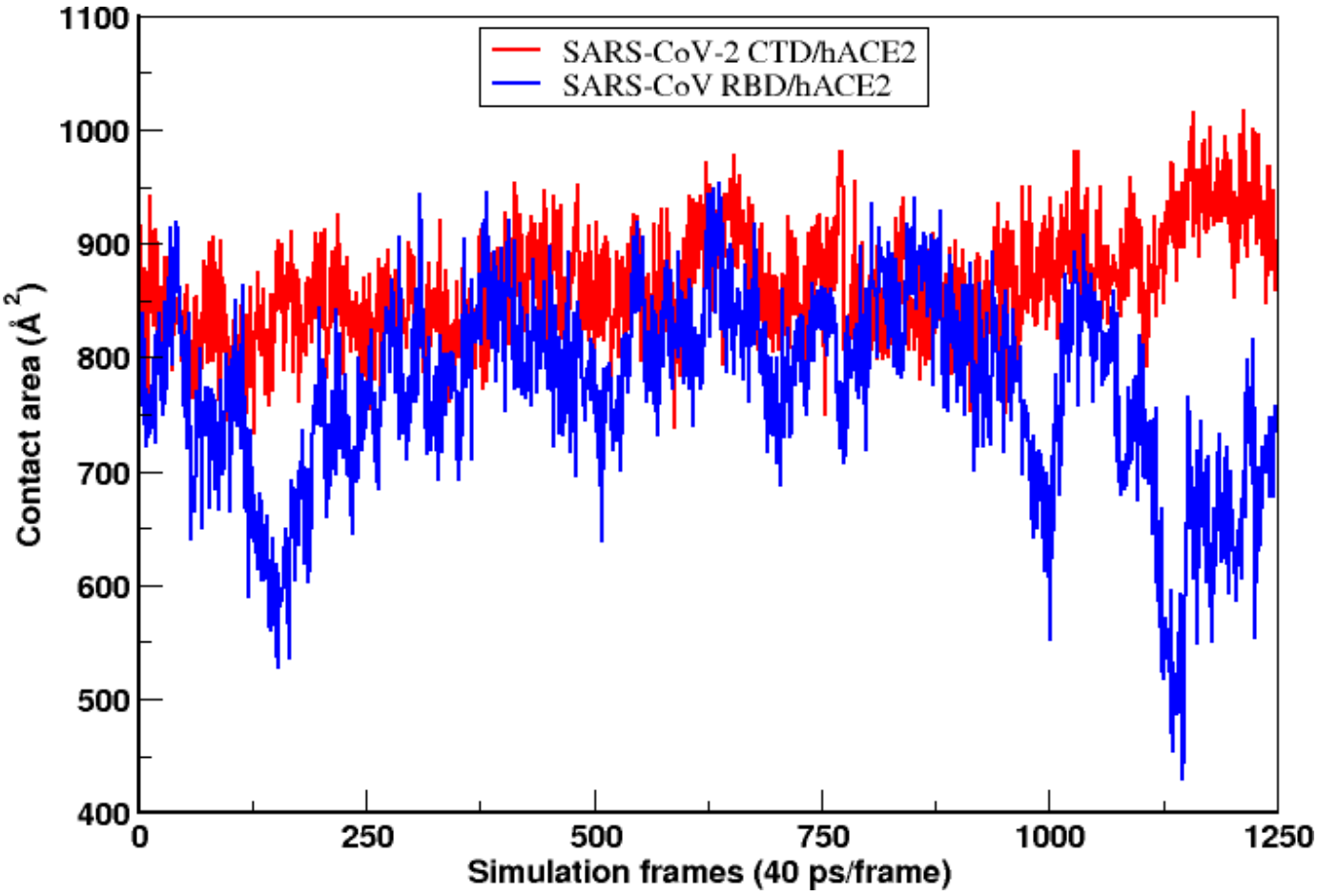


Figure 2

Comparison of contact surface area at the interface of SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2.

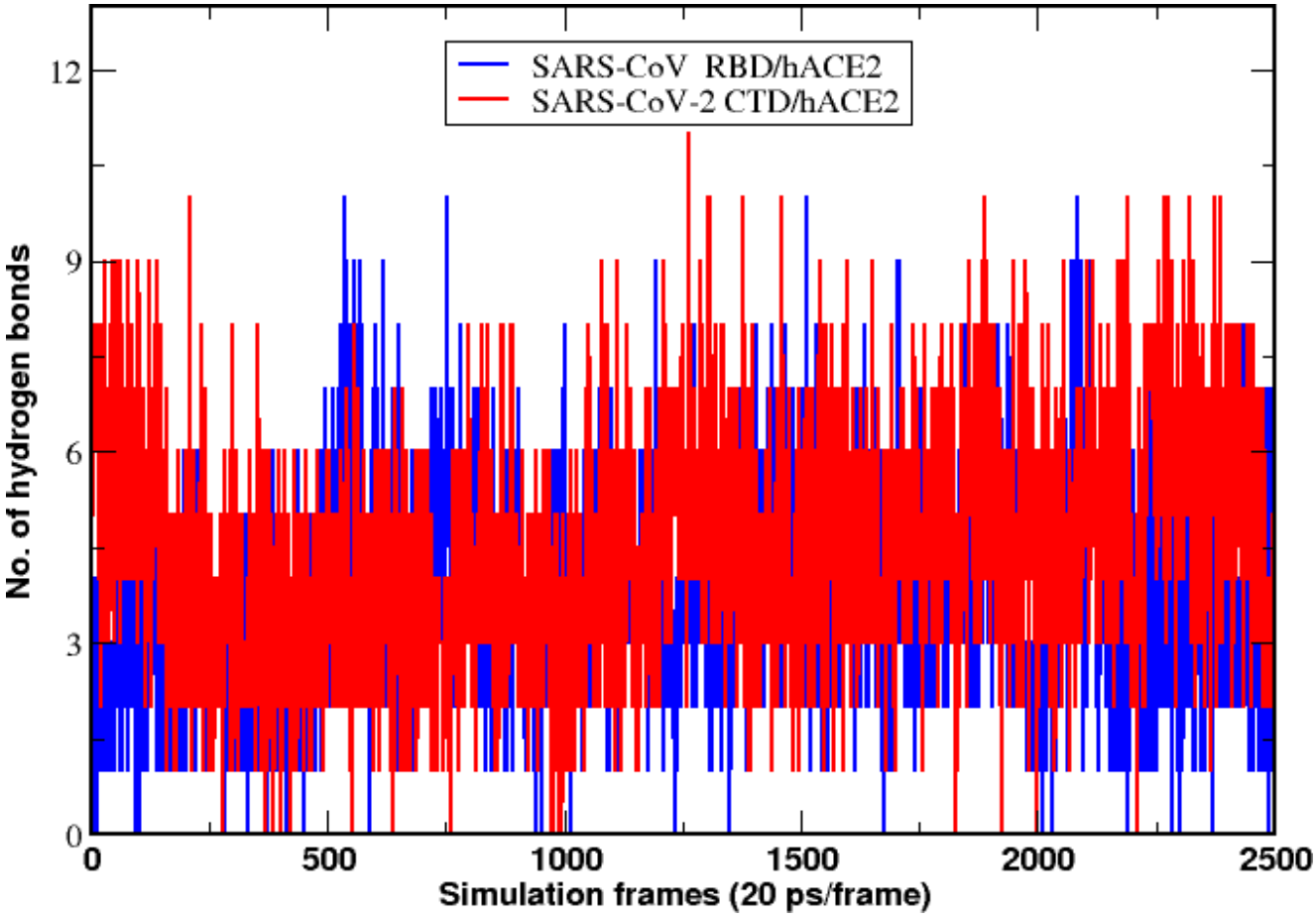


Figure 3

Comparison of time evolution of number of hydrogen bonds at the interfaces of SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2.

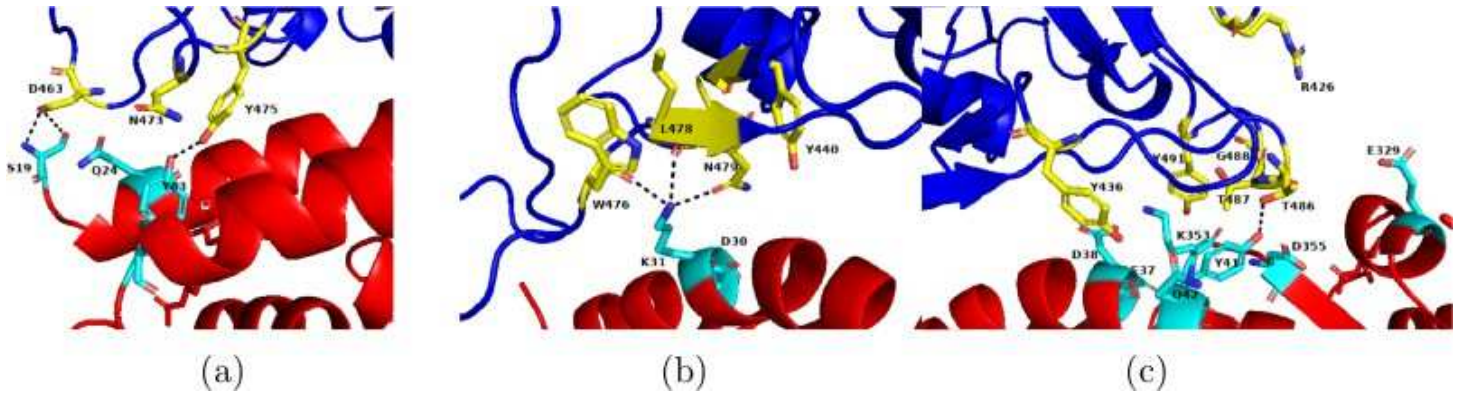


Figure 4

Details of hydrogen bondings at the three key contact sites on binding interface between SARS-CoV RBD and hACE2 receptor. hACE2 receptor is shown in red and SARS-CoV RBD is shown in blue color. The contact residues which can form potential hydrogen bonds are shown as sticks & labelled by 1 residue letter (for 0th frame).

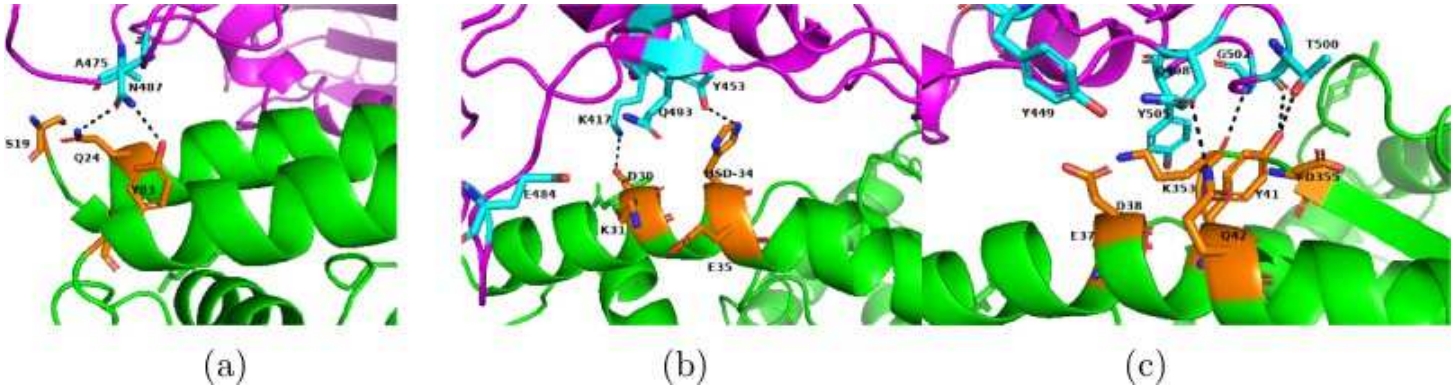


Figure 5

Details of hydrogen bondings at the three key contact sites on binding interface between SARS-CoV-2 CTD and hACE2 receptor. hACE2 receptor is shown in green and SARS-CoV-2 CTD is shown in magenta color. The contact residues which can form potential hydrogen bonds are shown as sticks & labelled by 1 residue letter (for 0th frame).

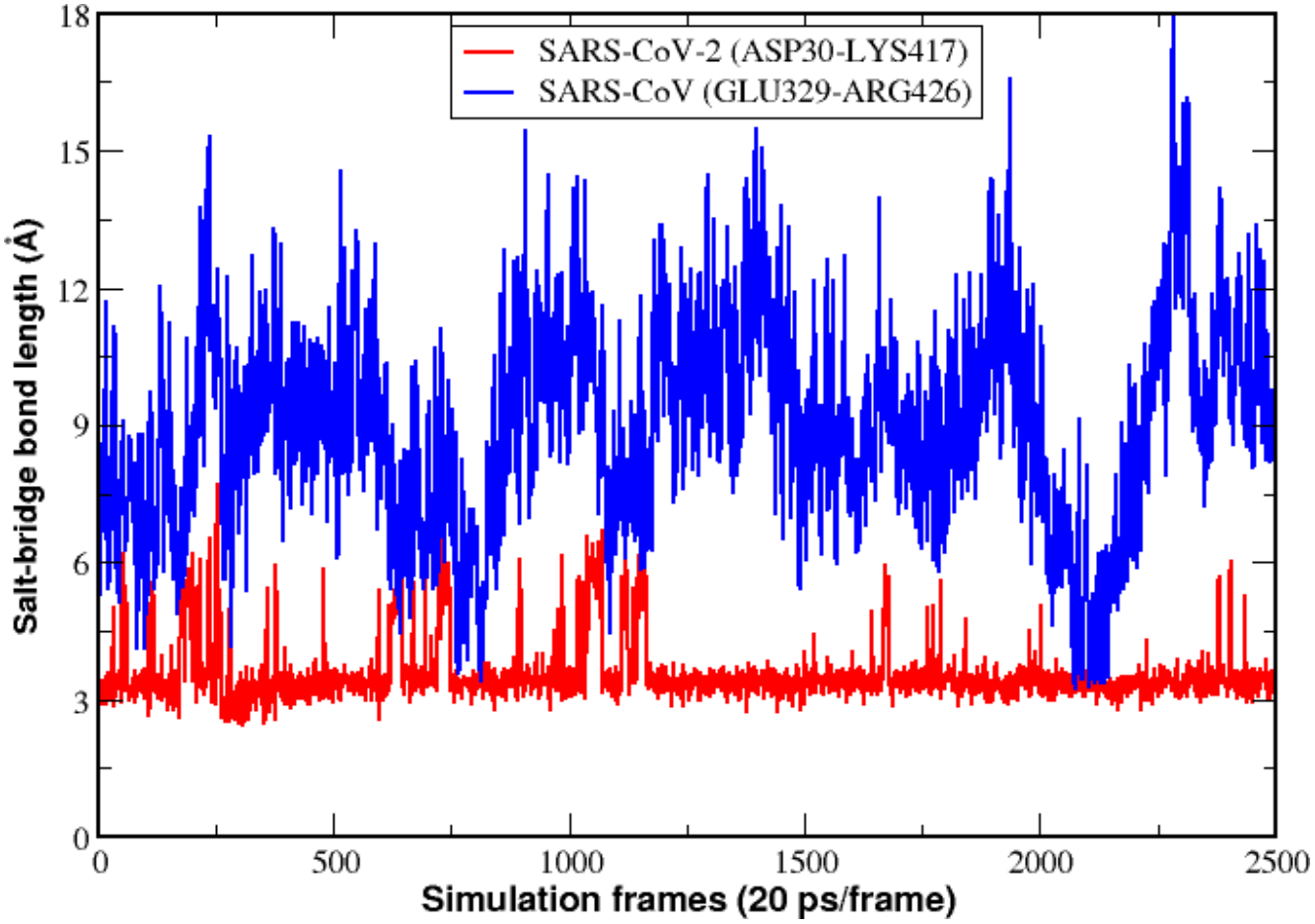


Figure 6

Comparison of time evolution of salt-bridge bond length at the interface of SARS-CoV RBD/hACE2 and SARS-CoV-2 CTD/hACE2.

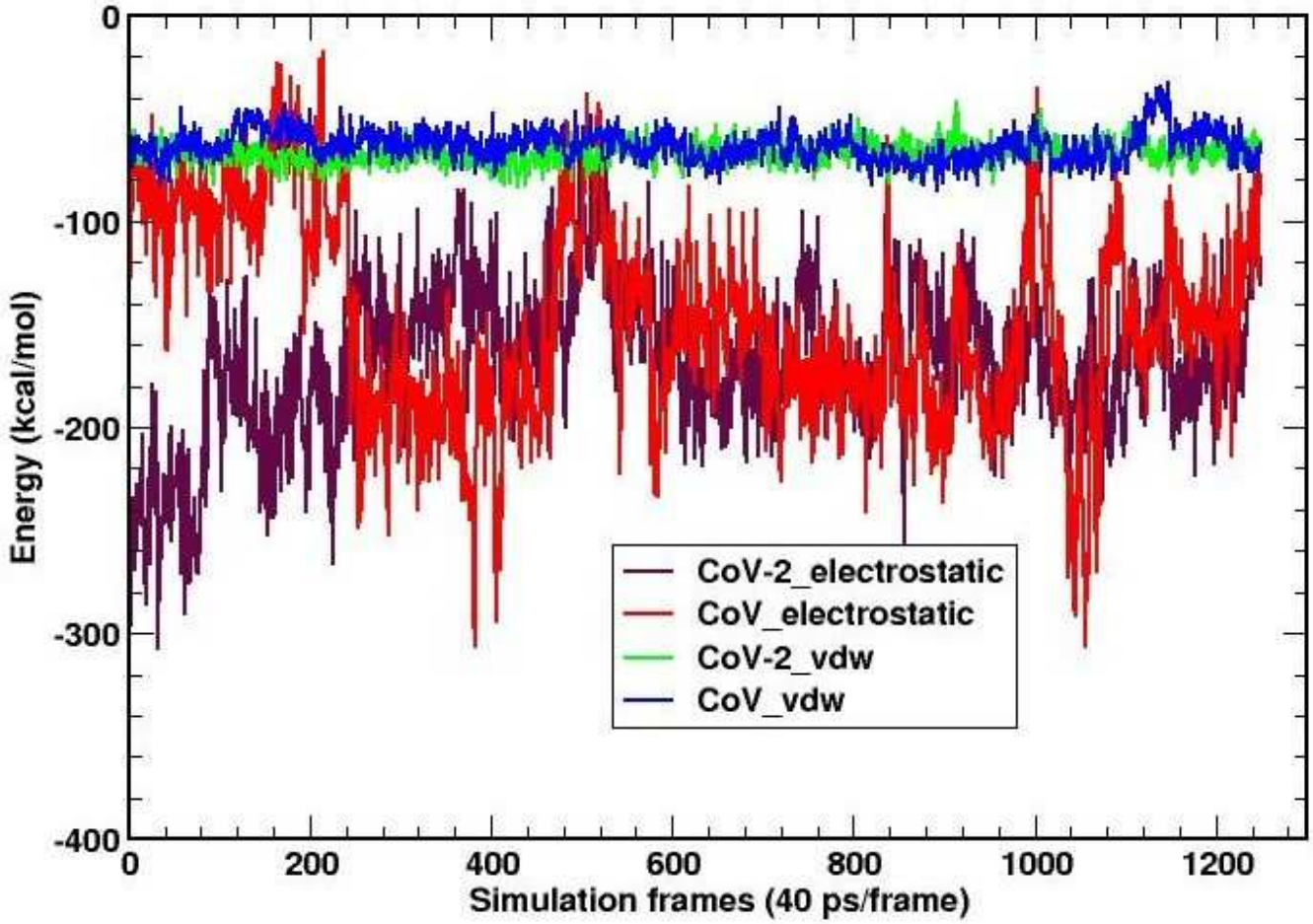


Figure 7

Comparison of electrostatic and vdw interactions of SARS-CoV-2 CTD and SARS-CoV RBD with hACE2.

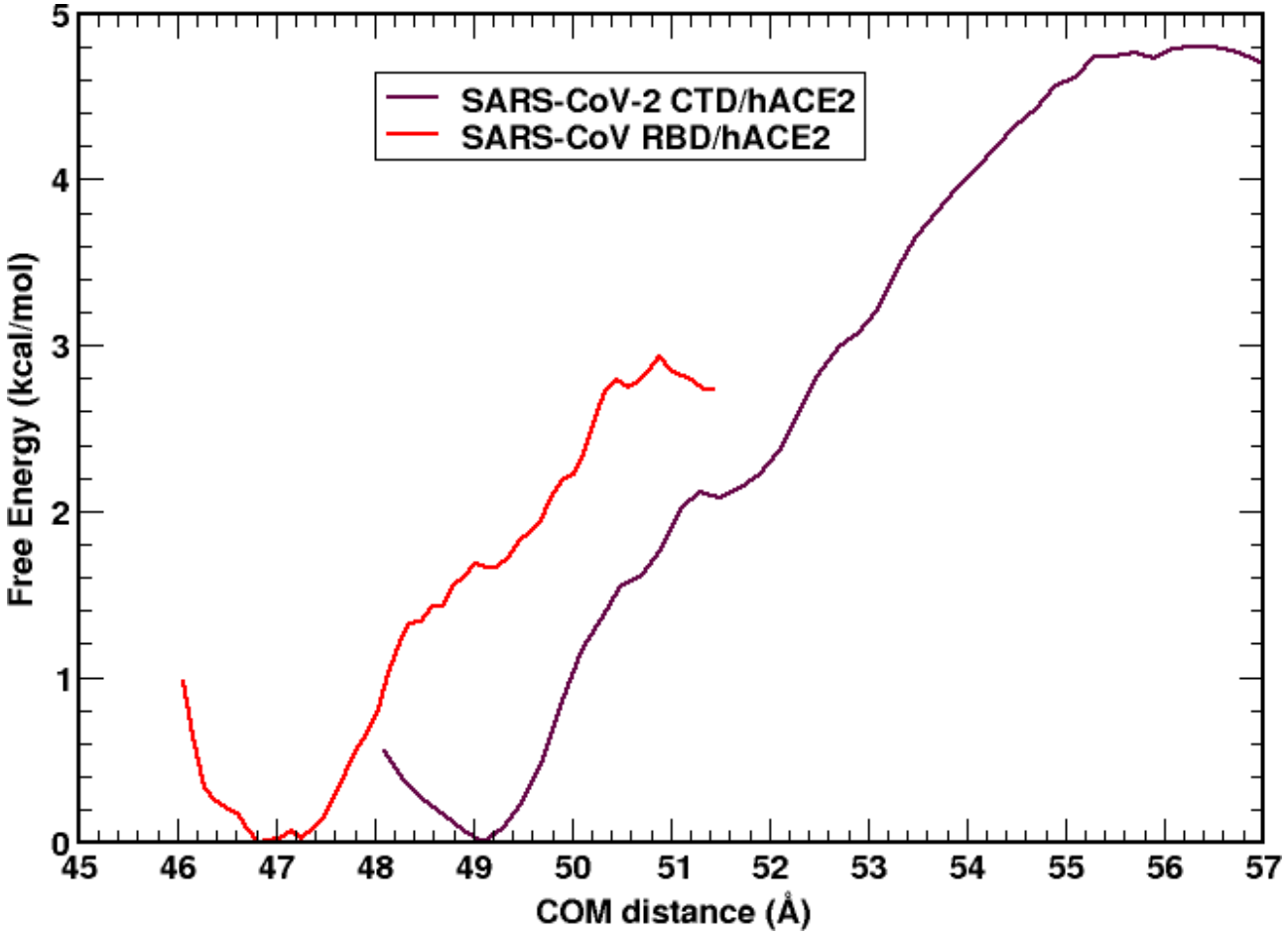


Figure 8

Free energy curve for SARS-CoV-2 CTD/hACE2 and SARS-CoV RBD/hACE2 during the translation of spikes CTD/RBD from hACE2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- CORONASupplementarymaterials.pdf

BIBECHANA
ISSN 2091-0762 (Print), 2382-5340 (Online)
Journal homepage: http://nepjol.info/index.php/BIBECHANA
Publisher: Department of Physics, Mahendra Morang A.M. Campus, TU, Biratnagar, Nepal

Thermal properties of normal and sickled hemoglobin protein

Jhulan Powrel[1,2], Narayan P. Adhikari[2]*
[1]Department of Physics, Butwal Multiple Campus, Tribhuvan University, Nepal
[2]Central Department of Physics, Tribhuvan University, Kirtipur Kathmandu, Nepal
*Email: narayan.adhikari@cdp.tu.edu.np

Article Information:
Received: June 29, 2020
Accepted: August 2, 2020

Keywords:
Thermodynamic properties
Sickle cell
Molecular dynamics
Hemoglobin protein

ABSTRACT
Thermodynamic properties of sickled and normal hemoglobin protein are considered within the framework of classical molecular dynamics. Here we have studied the specific heat capacity and RMSD (Root Mean Square Deviation) of both types of hemoglobin protein. Our investigation reveals that the specific heat capacity and RMSD for oxygenated hemoglobin protein is higher than those of de-oxygenated sickle hemoglobin protein. It is also observed that the specific heat capacity and RMSD values of sickle hemoglobin protein decrease with rise in temperature.

DOI: https://doi.org/10.3126/bibechana.v18i1.29750

1. Introduction

Hemoglobin is the blood component consisting of heme and globin, where heme is the prothestic group containing iron and globin is the protein compound. It is responsible for the transportation of oxygen throughout the human body. Heme is synthesized by mitochondrion which is fixed with iron. Heme is protected by the surrounded globin protein. Each single hemoglobin molecule has two globin chains, one globin chain is α and the other is β. Two hemoglobin molecules combine to produce a functional tetramer. In inherited hemoglobin the amino acid sequence is changed because of the incorrect DNA code (HbS). Such a cell having abnormality is a sickle cell. Sickle cell disease is the genetic disorder in which sudden change of hemoglobin protein occurs. The glutamic acid of both the β chain in the sixth position of normal hemoglobin protein is replaced by valine in the sickle hemoglobin which is shown in Figure 1. As a result RBC changes from flexible bi-concave to strong elongated shape. Thus by carrying less oxygen, the hemoglobin becomes fibrous aggregations giving the entire red blood cell a sickle shape. Sickled hemoglobin protein (HbS) damages the sickle erythrocyte and causes the increase

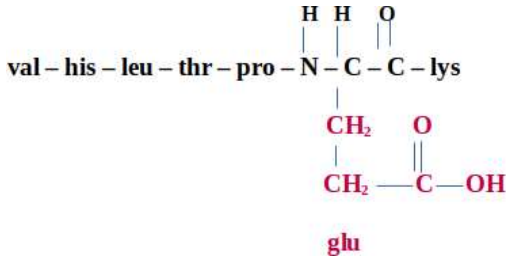
in the density, reduction in the deformability and increase in the adhesivity which shortens the life span [1-3]. The polymerization of deoxygenated hemoglobin (HbS) is the primary and indispensable event in the molecular pathogenesis of sickle cell disease. There are many symptoms like the attacks of pain, sickle cell anemia, swelling in the hands and feet, bacterial infections and stroke. Long-term pain may develop as people get older because of sickle cell anemia. On the other hand, heterozygote for the sickle gene are relatively protected against the danger of dying by malaria, as now firmly established through a number of clinical field studies from different parts of Africa and Asia(Terai region of Nepal)[4].

Figure 2(a) shows the glutamate 6 of α chain in hemoglobin protein which is undergoes a mutation to valine. This mutation changes the charge on the surface of hemoglobin. The mutant protein is called sickle cell hemoglobin (HbS). Valine residue 6 of the β chain of deoxy-HbS lies on the surface of the protein. This hydrophobic residue, present in each of the β chain, forms a hydrophobic contact in the neighboring β chain of another hemoglobin molecule [1, 6]. Figure 2(b) shows the mutation that replaces a glutamate residue by a valine residue which reduced the solubility of the protein.

Luzzatto L. studied the structural abnormality of sickle hemoglobin (HbS) and noticed as this was the first time that a single amino acid replacement in a protein causes a serious disease [7]. A group of scientists, Stradner and co-worker studied molecular dynamical simulations [1] and developed the concept of the study from soft matter physics to complex protein mixtures [8].A potential antisickling drug is still to be discovered [9] as per pathophysiological study in the last 100 years of sickle cell disease. Another group, Xianqiao Wang and Rodney D. developed a virtual mechanical model for determining the mechanical and thermal behavior of different soft matter by using the steps of optimization of energy, making of equilibration in the system and production run [10]. An electro-chemistry based study designed a technique to control and monitor the polymerization of sickle-cell hemoglobin (HbS) by looking the change in turbidity during the depletion of oxygen in a small volume of custom-built thin-layer electrochemical cell. The cell allowed the investigation of HbS polymerization as a function of HbS concentration, temperature and solution pH [11].

Due to bad impact of sickle cell disease, this research work got attention on the study of thermal and transport property of sickle hemoglobin and normal hemoglobin protein. Literature review also reveals that the mechanical, thermal and transport properties of normal hemoglobin and sickle hemoglobin protein around human body temperature (310 K) are not studied yet.

The paper is organized with methods and methodology part in section 2, computational simulation part in section 3, results with discussion in section 4 and finally the conclusions at the last.



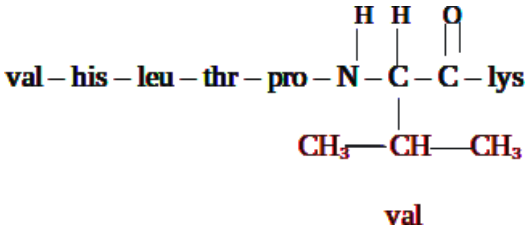


Fig. 1: Normal (glu) and sickle cell (val) hemoglobin [5].

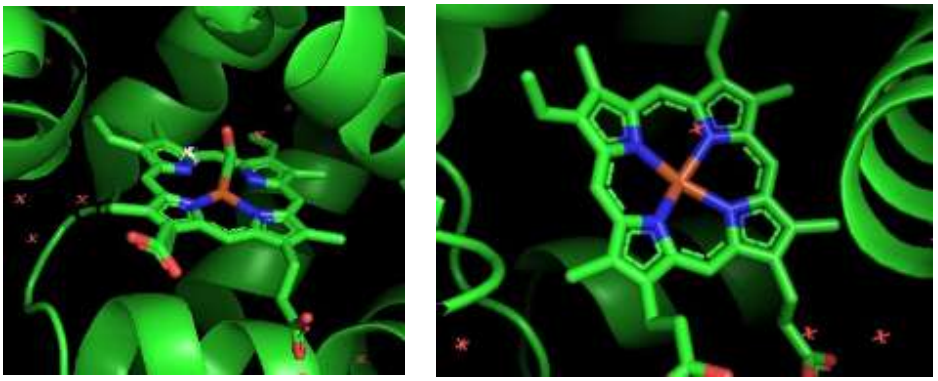


Fig. 2: (a) Normal hemoglobin with glutamic acid in 6th position of beta chain (left) and (b) Sickle hemoglobin protein with valine in the 6*th* position of beta chain (right).

2. Methods and Methodology

Theory

In molecular dynamics (MD), a hemoglobin protein molecule is considered inside the water for simulation. Protein is the combination of a number of atoms which behaves as pair-wise interaction in between two atoms in the molecule [12]. Each pair of atoms is interacted by bonded and non-bonded potentials. The molecular dynamic simulation calculates the motion of the atoms in molecules using Newtonian dynamics. In standard molecular dynamics simulations the forces may come from classical inter atomic potentials [13]. The non-bonded van der Waals and electrostatic are the interactive forces among the particles in the hemoglobin molecules. The total potential is obtain as,

$$V_{total} = [V_{bonded}] + [V_{nonbonded}] \quad (1)$$

Or,

$$V_{total} = [V_{bond} + V_{angle} + V_{improper} + V_{proper}] + [V_{LJ} + V_{coloumb}] \quad (2)$$

Now the bond stretching potential is given by,

$$V_{bond}(r_{ij}) = \frac{1}{2} K^{b}_{ij}(r_{ij} - b_{ij})^2. \quad (3)$$

where K^{b}_{ij} force is constant, r_{ij} is separation between the two bounded particles and b_{ij} is the equilibrium bond length. Again the bond angle potential is written as,

$$V_{angle}(\theta_{ij}) = \frac{1}{2} K^{\theta}_{ijk}(\theta_{ijk} - \theta^{0}_{ijk})^2 \quad (4)$$

where θ_{ijk}, θ^{0}_{ijk} and K^{θ}_{ijk} are the bond angle, equilibrium bond angle and the force constant respectively. Now the improper dihedral potential energy is,

$$V_{improper}(\xi_{ijkl}) = \frac{1}{2} K_{\xi}(\xi_{ijkl} - \xi_0)^2 \quad (5)$$

where ξ_0 is the equilibrium improper angle and K_{ξ} is the constant. Similarly the dihedral potential is,

$$V_{proper}(\phi_{ijkl}) = \frac{1}{2} K_{\phi}(1 + \cos(n\phi_{ijkl} - \phi_s)) \quad (6)$$

where ϕ_{ijkl} is the proper dihedral angle, ϕ_s is an angle at which the potential is maximum, n is the multiplicity and K_{ϕ} is the barrier height.

Although Lennard-Jones potential(L-J) is not suitable for system where strong localized bonds may form (as in covalent system) or there is a delocalized electron sea where the ions remain (as in metals), the L-J potential helped us to understand basic points in many areas of condensed matter physics. The L-J potential for interaction between a pair of atom is given by [14],

$$V_{Lj} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \quad (7)$$

where ε is the depth of the attractive well, σ refers to the inter-particle distance and defines the strength of the interaction where the potential changes sign and r defines the length scale. The Coulomb electrostatic interaction between two point charges q_i and q_j is defined by Coulomb potential as,

$$V_{Coloumb} = \frac{q_i q_j}{4\pi\varepsilon_m r^2} \quad (8)$$

Root Mean Square Deviation (RMSD):

RMSD is the level of equilibration and gives the deviation of the molecules from the defined position in space. It also gives the idea of occurrence of the reaction in chemical bonding. The larger the deviation from mean position less will be the stability of the system. RMSD values are calculated for all atoms of the hemoglobin protein backbone for the entire protein and for the protein excluding the last five residues[15] , which is the numerical measure of the structural difference of two states given by,

$$RMSD(t_j) = \sqrt{\frac{\sum_{\alpha=1}^{N_\alpha}\left(\vec{r}_\alpha(t_j) - \left\langle \vec{r}_\alpha(t_j)\right\rangle\right)^2}{N_t}} \quad (9)$$

With

$$\left\langle \vec{r}_\alpha(t_j)\right\rangle = \frac{1}{N_t}\sum_{j=1}^{N_t}\vec{r}_\alpha(t_j) \quad (10)$$

where N_α is the number of atoms whose positions are compared, N_t is the number of time steps over which atomic positions are compared, $\vec{r}_\alpha(t_j)$ is the position of atom α at time t_j and $\left\langle \vec{r}_\alpha(t_j)\right\rangle$ is the average value of the position of atom to which the positions $\vec{r}_\alpha(t_j)$ are being compared[15].

Specific heat Capacity:

The specific heat capacity is the amount of heat required for raising the temperature of an object per degree of temperature increase per unit mass. Thus in canonical ensemble the specific heat capacity is [16],

$$c_v = \frac{\left\langle E_B{}^2\right\rangle - \left\langle E_B\right\rangle^2}{k_B T^2} \quad (11)$$

where $\left\langle E_B\right\rangle^2$ is square of the average of the total energies and $\left\langle E_B{}^2\right\rangle$ is the average of the square of the total energy and k_B the Boltzmann constant.

3. Computational details

The proposed research is designed for working in theoretical computational method using molecular dynamics. Hemoglobin protein of sickle and normal cell are obtained from the RCSB (Research Collaboratory for Structural Bioinformatics).org in PDB (Protein Data Bank) form and then simulation is completed using NAMD (NAno scale Molecular Dynamics) and VMD (Visual Molecular Dynamics) software by following the steps of structure making, energy minimization and making of equilibration in the system. After performing various simulation steps the output data are used for the calculations of RMSD and specific heat capacity of both of the hemoglobin protein. For this simulation six systems are designed with four molecules which are then solvated with water using charmm force field (CHARMM 36) under following parameters: A system of oxygenated hemoglobin protein (a3n) with 4546(without water) atoms under the interaction of the bond potential and non-

bonded potential inside a box with cut off distance of 12.0 Å, pressure of 1 bar and temperature 310 K. A system of oxygenated hemoglobin protein(e83) with 19167(without water is 4668) atoms under the interaction of the bond potential and non-bonded potential inside a box with cut off distance of 12.0 Å, pressure of 1 bar and temperature 310 K. A system of oxygenated hemoglobin protein(uvx) with 1863 atoms under the interaction of the bond potential and non-bonded potential inside a box with cut off distance of 12.0 Å, pressure of 1 bar and temperature 310 K. Similarly, for other three systems of sickle hemoglobin protein(HbS) there are 9112 atoms in each and interacting at the same potential as before with pressure at 1 bar, cut off distance of 12.0 Å and at temperatures of 305 K, 310 K and 315 K. After equilibration of all the above six system they are energetically minimized at their corresponding constant temperature which provides their RMSD values. At last systems are further modified for NVE and NVT simulation run in which we use the temperature coupling feature of NAMD to set the temperature of the molecules in the outer layer of the sphere to 300 K, while the rest of the bob is set to temperature of 330 K.

4. Result and Discussion

Present work was to study the RMSD as well as specific heat capacity of normal and sickle hemoglobin protein. The RMSD was calculated using the equation (9).

We plotted the RMSD of hemoglobin protein against the time steps as follows; where Figure 3(a) and 3(b) indicates the graph of root mean square deviations of oxygenated sickle hemoglobin protein and de-oxygenated hemoglobin protein at temperature of 310 K and obtained the RMSD as 0.8 Å and 1.25 Å respectively. They were found to be increased initially with small fluctuation and then remain constant for both at various values, i.e. the RMSD of deoxygenated hemoglobin found to be higher than of oxygenated sickle hemoglobin protein.

RMSD plot of other system of hemoglobin protein were also plotted where Figure 4(a) and 4(b) indicates the root mean square deviations of oxygenated hemoglobin protein and sickle hemoglobin protein at temperature of 310 K. For which the RMSD found to be increased initially with small fluctuation and then remain constant for both at various values of 1.3 Å and 0.75 Å respectively, i.e. the RMSD of oxygenated hemoglobin is higher than of sickle hemoglobin protein.

Similarly, RMSD for sickle hemoglobin protein at various temperatures were studied, where Figure 5(a) and 5(b) indicating the root mean square deviations of sickle hemoglobin protein at 305 K and 315 K respectively. They are increased initially with small fluctuation and then remain constant for both at various values of 1.5 Å and 1.4 Å respectively i.e. the RMSD of sickle hemoglobin protein at lower temperature is higher than of sickle hemoglobin protein at higher temperature. The above Figure 3(a), 3(b), 4(a) and 5(a), are the RMSD graph at 310 K for carbon monoxy-hemoglobin(s), de-oxygenated hemoglobin protein, truncated heme-ligand hemoglobin and sickle hemoglobin protein respectively. Correspondingly their RMSD were estimated as 0.8 Å, 1.2 Å, 1.3 Å and 0.75 Å, i.e. the RMSD of sickle hemoglobin is least(0.75 Å) and of truncated ligand hemoglobin is maximum(1.3 Å). Thus, deoxygenated sickle hemoglobin with fibrous nature shows least deviation from the mean position, whereas of truncated ligand hemoglobin with large no of hydrogen bond and hydrophobic potential move faster with more RMSD value. Similarly normal oxygenated and deoxygenated hemoglobin

have their RMSD 1.3 Å and 1.25 Å respectively which are higher than both the oxygenated (0.8 Å) and deoxygenated sickle hemoglobin (0.75 Å). Again at temperatures of 305 K, 310 K and 315 K (Figure 4(b), 5(a) and Figure.5 (b)) of de-oxygenated sickle hemoglobin (HbS) [17] the RMSD are 1.5 Å, 0.75 Å and 1.4 Å respectively.

Figure 6(a) and 6(b) are energy fluctuation and energy distribution graph during NVE and NVT run. They gives the energy deviation from the mean [Fig 6(a)] and energy distribution pattern [Fig 6(b)] i.e. more the energy more is the velocity as well as temperature of the system of particles and hence system of protein behaving as thermometer. Computationally by using NAMD code the average of square of the total energies and square of the average total energy of de-oxygenated hemoglobin protein(e83) were obtain as 6722047.08 $kcalmol^{-1}$ and 6714152.23 $kcalmol^{-1}$ respectively. On using these values the specific heat capacity of de-oxygenated hemoglobin protein was estimated as 2903 $Jkg^{-1}C^{-1}$ (Table 1). Now for the calculation of specific heat capacity (C_v), we have used the following theoretical and simulation parameters under the assumption of classical molecular dynamics: k_B=0.00198657 kcal /mol-K, σ ~ 8 K, T=317 K, J=1.43846 $\times 10^{20}$ kcal/mol.

The specific heat capacity of normal and sickle hemoglobin protein were calculated by using the equation (11) and are tabulated in Table 1 and Table 2 for all the six systems.

The average specific heat capacity of sickle hemoglobin ($\frac{2527+1966+1227}{3} = 1906.67$ J/kg-C [Table 1]) was less than of normal hemoglobin's (3587 J/kg-C, 2903 J/kg-C and 3204 J/kg-C) [Table:2].

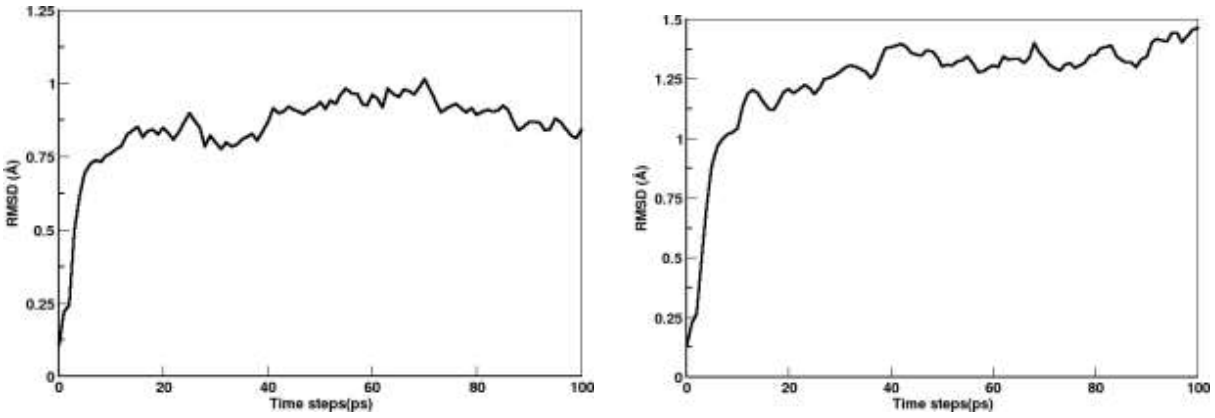


Fig. 3: (a) RMSD graph of oxygenated sickle hemoglobin protein (left). (b) RMSD graph of deoxygenated hemoglobin protein (right).

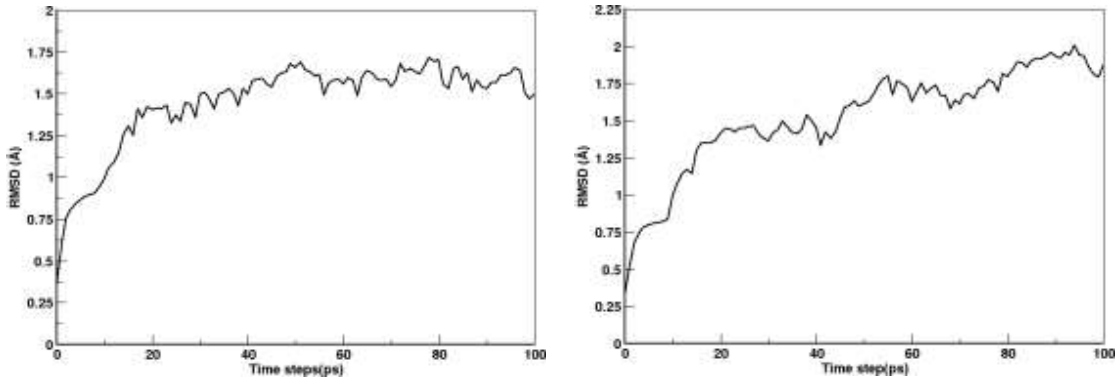


Fig. 4: (a) RMSD graph of oxygenated hemoglobin protein (left). (b) RMSD graph of sickle hemoglobin protein (right).

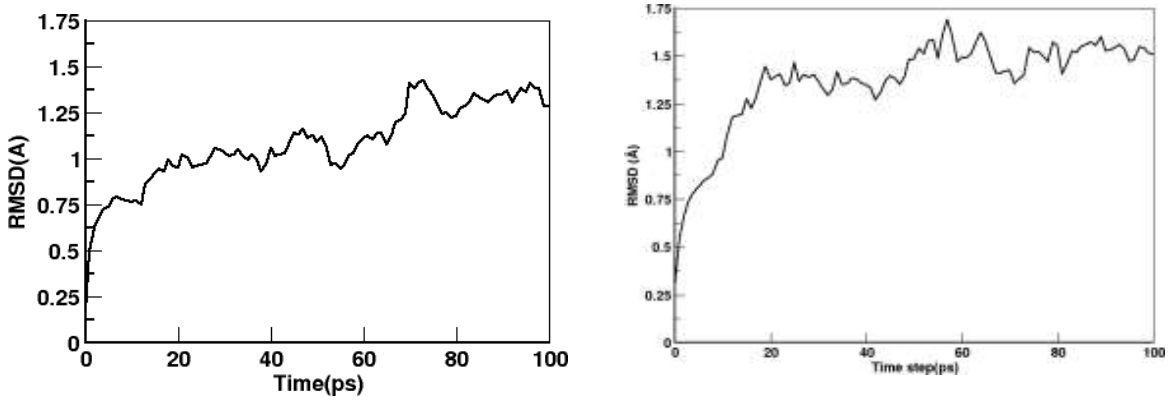


Fig. 5: (a) RMSD graph of sickle hemoglobin protein at 305 K (left). (b) RMSD graph of sickle hemoglobin protein at 315 K (right).

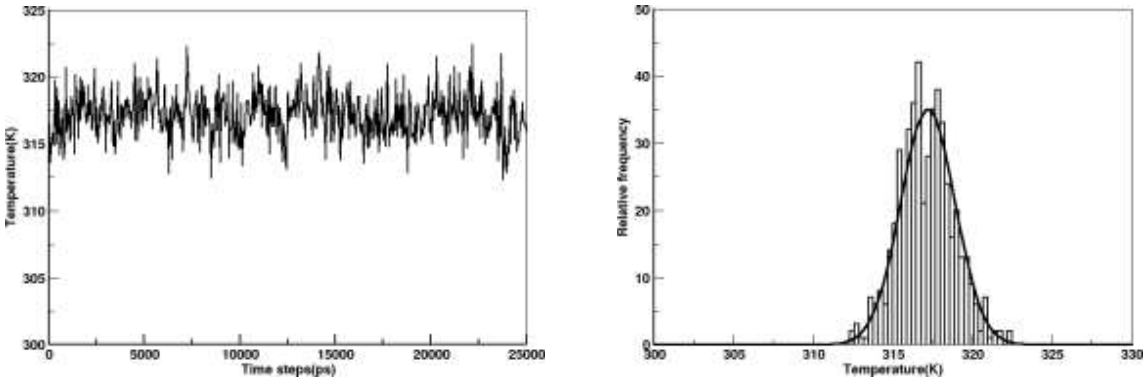


Fig. 6: (a) Energy flocculation of de-oxygenated hemoglobin protein (b) Maxwell energy distribution for de-oxygenated hemoglobin protein.

Table 1: For normal and sickle cell hemoglobin protein at 310 K of temperature.

For	RMSD(Å)	C_v(Jkg^{-1}C^{-1})
Oxygenated sickle hemoglobin protein)	0.8	3587
De-oxygenated hemoglobin protein	1.25	2903
Oxygenated hemoglobin protein	1.3	3204
Sickle hemoglobin protein(2hbs)	0.75	1966

Table 2: For sickle cell hemoglobin protein (2hbs) at different temperature.

For	Temp.(K)	RMSD(Å)	C_v(Jkg^{-1}C^{-1})
Sickle hemoglobin protein(2hbs)	305	1.5	2527
Sickle hemoglobin protein(2hbs)	310	0.75	1966
Sickle hemoglobin protein(2hbs)	315	1.4	1227

5. Conclusion and Concluding Remarks

The RMSD of normal and sickle hemoglobin protein was calculated within 0.75 Å to 1.5 Å which indicates that the proposed system is in the valid range(< 4 Å) for molecular dynamical study. This study also reveals that the HbS at 305 K is most unstable (RMSD=1.5 Å) among the considered system. The RMSD and specific heat capacity (C_v) of de-oxygenated hemoglobin protein at 310 K are estimated as 1.25 Å and 2903 J/kg-C [Table 1] respectively. It also obvious that the de-oxygenated hemoglobin protein (average value=1906.67J/kg-C) has lower specific heat capacity than oxygenated (first three C_v of Table 1). It shows that at higher (315 K) temperature above 310 K the RMSD of sickle hemoglobin protein is increased due to less polymerization (sickling) of protein but increased in RMSD at lower temperature (305 K) is still not understood.

We found essentiality of additional research work on the free energy calculation along with the mechanical and transport properties of sickle hemoglobin protein for its depth knowledge.

Acknowledgments

JP acknowledges UGC, Nepal for providing financial support (Grant No. S&T-13-75/76) and Butwal Multiple Campus, Butwal (Tribhuvan University) for providing Ph. D. leave. This work is also under the grant of UGC, Nepal for which NPA acknowledges the UGC Award no. CRG-73/74-S&T-01.

References

[1] A. Stradner et al, New insight into cataract formation â enhanced stability through mutual attraction, Physics Department and Fribourg Center for Nanomaterials, University of Fribourg, CH-1700 Fribourg, Switzerland (2007).

[2] W. A. Eaton and J. Hofrichter, Sickle cell hemoglobin polymerization, Advances in Protein Chemistry 40 (1990) 63-279. http://doi.org/10.1016/s0065-3233(08)60287-9.

[3] F. Ferrone, R. L. Nagel, Sickle hemoglobin polymerization. In: Disorders of hemoglobin: Genetics, pathophysiology, clinical management, Steinberg MH, For-get BG, Higgs D, Nagel RL (Eds), Cambridge University Press (2000).

[4] Luzzatto, Lucio., Sickle cell anaemia and malaria, Mediterranean Journal of Hematology and Infectious Diseases (4)1 (2012) e2012065. http://doi.org/10.4084/MJHID.2012.065.

[5] D. C. Rees, T. N. Williams, M. T. Gladwin, Sickle-cell disease, The Lancet 376(9757) (2010) 2018-2031.

[6] J. Robert, J. Ouellette, David Rawn, Organic Chemistry (Second Edition) Structure, Mechanism, Synthesis (2018) pp. 929-971.

[7] W. D. Callister, Materials Science and Engeenering, An Introduction, John Wiley Sons.Inc. US, (2007).

[8] Retrieved Sept. 22th, 2018 from www.google.com/Sickle Cell Anemia.

[9] Mahran et al., Beni-Suef University Journal of Basic and Applied Sciences, Springer open, Egypt (2019). https : //doi.org/10.1186/s43088 − 019 − 0016 − x

[10] Wang Xianqiao and D. Rodne, Avereabtt Molecular dynamics simulations indicate that deoxyhemoglobin, oxyhemoglobin, carboxy-hemoglobin, and glycated haemoglobin hemoglobin under compression and shear exhibit an anisotropic mechanical behavior, Sumith Yesudasan ORCID Icon (2017)1417-1429. https://doi.org/10.1080/07391102.2017.1323674.

[11] Iqbal, Zeshan et al., Investigation of sickle-cell polymerisation under electrochemical control, Chemphyschem: A European Journal of Chemical Physics and Physical Chemistry 14 (10) (2013) 2143-8. https://doi.org/10.1002/cphc.201300203.

[12] G. R. Serjeant, Sickle-cell disease. The Lancet 350(9079) (1997) 725-730

[13] M. P. Allen and D. J. Tildesley., Computer Simulation of Liquids, Oxford University Press, New York (1987).

[14] D. Frenkel, B. Smit, Understanding Molecular Simulation From Algorithms to Applications 2nd ed; Academic Press; San Diago, USA (2002).

[15] C. James Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry 26 (2005)1781-1802. www.ks.uiuc.edu/Training/Tutorials(RetrivedonJune10th2020

[16] K. Huang, Statistical Mechanics, John Wiley & Sons, New York (1987).

[17] J. C. Phillips et al., Scalable molecular dynamics with NAMD, J. Comp. Chem. 26 (2005) 1781-1802.

Butwal Campus Journal, Vol. 5, No. 1: 143-159, July 2022
Research Management Cell, Butwal Multiple Campus, Tribhuvan University, Nepal
DOI: https://doi.org/10.3126/bcj.v5i1.50199

COMPUTATIONAL ESTIMATION OF FREE ENERGY USING FORTRAN CODE

Jhulan Powrel ,

Associate Professor

Butwal Multiple Campus, Department of Physics, Tribhuvan University, Nepal

Article History: *Received 14 May 2022; Reviewed 09 June 2022; Revised 15 July 2022; Accepted 25 July 2022*

ABSTRACT

Theoretical physics based on several computational software. The code for theoretical calculations of free energy is essential in many of the software. One of the codes frequently used in molecular dynamics is FORTRAN (Formula Transformation) code. The free energy for thermal and conformational are found to be $0.2\ \mathrm{KJ/mol}$ and $0.2\ \mathrm{KJ/mol}$ respectively. The average of the total free energy of the theoretical system is found to be $0.40\ \mathrm{KJ/mol}$. The free energy of a system of theoretical molecule of butane is $(-2.75 \pm 0.16)\ \mathrm{KJ/mol}$. The ranges of theoretical and experimental values of free energy are found to be in considerable range. So, the FORTRAN code can also be used in estimating the free energy of the system of molecules.

Keywords: Computational, FORTRAN code, Probability density, Conformational energy, Thermodynamics energy

INTRODUCTION

The free energy is the energy associated with a chemical reaction that can be used to do work. The free energy of a system is the sum of its enthalpy (H) and the product of the temperature (Kelvin) and the entropy (S) of the system i.e., $\mathrm{G = H - TS}$. It is also the change in the enthalpy (H) of the system minus the product of the temperature (Kelvin) and the change in the entropy (S) of the system:

$$\Delta G = \Delta H - T\Delta S$$

The change in free energy that occurs when a compound is formed form its elements in their most thermodynamically stable states at standard-state conditions. In other words, it is the difference between the free energy of a substance and the free energies of its elements in their most thermodynamically stable states at standard-state conditions. The standard-state free energy of reaction can be calculated from the standard-state free energies of formation as well. It is the sum of the free energies of formation of the products minus the sum of the free energies of formation of the reactants. The free energy is the thermodynamic free energy and is the useful work that may be extracted from a closed system. It is the function of energy of system (enthalpy), entropy and temperature of the system (Kittel & McEuen, 1976). It may be zero / positive or negative. Free energy calculation results reproduced for different model of water in experimental work are in well agreement with computational simulation. Free energies for different models are within 1.6 kJ mol^{-1} and entropies are within 3.6 JK^{-1} mol^{-1}. Approximately two-thirds of the entropy comes from translation, a third from rotation, and 5% from conformation. Intermolecular quantum effects for the classical oscillator lying within 0.5 to 0.7 KJmol^{-1}, which is higher than that in the quantum case. The method of thermodynamic integration, direct particle insertion, SMD and grand canonical Monte-Carlo are the different way for calculation of free energy. Free-energy calculations have a fundamental role in the understanding of many natural phenomena ranging from protein folding up to polymorphic transitions in solids. To this end, molecular dynamics (MD) has been extensively used together with a series of algorithms aimed at extending its capabilities far beyond those allowed by straightforward MD. Among some of the most popular enhanced sampling techniques are umbrella sampling (Stradner et al., 2007), Jarzynski equation based methods and umbrella sampling (Jarzynski, 1997). So umbrella sampling is introduced here.

Statistics of free energy

In the case of N particle system in 3D space the system is defined by the Hamiltonian which is the function of position and momentum given by,

$$H(p,q) = \sum_{1}^{N} \frac{P_i^2}{2m_i} + V(r_1, r_2, r_3, \dots\dots, r_N)$$

Thus,

$$F(N,V,T) = -K_B T * \ln \left[\frac{1}{h^{3N}} \iint \exp\left[-\frac{H(p,q)}{K_B T}\right] dpdq\right.$$

Here V-the volume of the system, T-absolute temperature of the system, K_B-Boltzmann constant and h-Planck's constant. As free energy is the change in the energy of the two corresponding state of a closed system hence it is expressed as,

$$\Delta F = F_B - F_A = -K_B T \ln \frac{N_B}{N_A}$$

Thermodynamic integration (TI)

As the state is changed the correspondingly interaction parameter is also changed in between $\lambda = 0$ and $\lambda = 1$, so the Hamiltonian of the system can be written as (DeVries & Hamill, 1995),

$$H \equiv H(p,q,\lambda)$$

Now for state A,

$$H_A(p,q) = H(p,q,\lambda_A),$$

and for state B,

$$H_B(p,q) = H(p,q,\lambda_B)$$

i.e. the Hamiltonian is function of λ, thus the equation for F gives,

$$\frac{dF(\lambda)}{d\lambda} = \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda$$

So the free emery difference of state A and state B of anmolecular system is ,

or, $F_{\lambda_B} - F_{\lambda_A} = \int_{\lambda_A}^{\lambda^B} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda,$

which is the thermodynamic integration (TI) formula. It gives the ensemble average, $\left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda$ from the simulation at a series of $\lambda' s$ between $\lambda_A = 0$ to $\lambda_B = 1$ (after integration). Now,

$$\text{or, } F_{\lambda_B} - F_{\lambda_A} = \sum_{1}^{\infty} \frac{dF(\lambda)}{d\lambda}(\lambda_B - \lambda_A)^n/n!$$

Umbrella sampling

The method of umbrella sampling is introduced by Torrie and Valleau in $1974-1977$. In this method blue moon ensemble approach along with some restrained factor of biasing potential instead of reaction coordinate is used. The biasing potential is the harmonic one called umbrella potential as,

$$W(f_1(r_1, r_2, \dots, r_N), S) = \frac{1}{2}K(f_1(r_1, r_2, \dots, r_N) - s)^2.$$

A total potential of $U(r) + W(f_1$,s) is incorporated in molecular dynamics simulation (Tuckerman, 2001). The U's are choosen at various intermediate points $s^{(1)}, \dots s^{(n)}$ in between initial $s^{(i)}$ and final $s^{(f)}$. Each of the reaction coordinate $\mathrm{q} = f_1(r)$ lies in between two end point values. Thus the new probability distribution function $\mathrm{P}(\mathrm{s}, s^k)$ gives the profile of the true free energy throughout the entire range of the reaction coordinate. An unbiasing techniques of weighted histogram analysis method (Yang & MacKerell Jr, 2015) is implemented in finding the biased probability distribution from each of the molecular dynamics,

$P(s, s^k) = \exp[\beta A_k] \int d^N r e^{-\beta\ (r)} e^{-\beta} W(f_1, s^k)\delta(f_1(r) - q)$,

Where, A_k the biased free energy and

$$e^{-\beta A_k} = \int d^N r e^{-\beta U(r)} e^{-\beta} W(f_1, s^k) = e^{-\beta A_0} \left\langle e^{-\beta W(f_1(r), s^k)} \right\rangle$$

This average is obtain by considering unbiased potential U(r) and

$$e^{-\beta A_0} = \int d^N r e^{-\beta U(r)}$$

which is the configuration partition function for unbiased case and $\exp(\beta A_k)$ is the normalization constant. Now the corresponding unbiased probability distribution function is,

$$P'(q) = e^{-\beta}(A_k - A_0)e^{\beta} W(q, s^k) P(q, s_k)$$

Hence the full probability distribution function is obtained in terms of other two is,

$$P''(q) = \Sigma_1^n C_k(q) P_k(q) = {}_1^n C_k(q)\{e^{-\beta}(A_k - A_0)e^{\beta} W(q, s^k) P(q, s_k)\}$$

Here $C_k(q)$ is optimised for accurate P''(q) and is $\Sigma C_k(q) = 1$. In oredr to find the coefficient $C_k(q)$ using WHAM (Yang & MacKerell Jr, 2015), let us write $H_k(q)$, biased histogram of each MD, then

$$P(q, s^k) \approx \frac{1}{n_k \Delta q},$$

Where, Δq the bin width and n_k the number of configuration sampled for k simulation. The statistical error for kth window is then,

$$\sigma_k^2 = \epsilon_k(q) H_k(q)/(n_k \Delta q)$$

This epsilon measure the deviation of numerically sampled and actual distribution in the kth window. Now the corresponding unbiased factor is,

$$\sigma'^2_k = \left\{e^{-2\beta}(A_k - A_0)e^{2\beta W(q,s^k)}\sigma_k^2\right\}$$

To minimize the total error let us introduce a multiplier λ as, in the error function,

$$\Sigma^2 = \sum_{k=1}^{n} C_k^2(q)\left\{e^{-2\beta}(A_k - A_0)e^{2\beta W(q,s^k)}\frac{\epsilon_k(q) H_k(q)}{n_k \Delta_k} - \lambda(\Sigma_{k=1}^n C_k(q) - 1)\right.$$

Taking derivation and then on solving in respect of λ

$$C_k(q) = \frac{\lambda n_k \Delta q}{2\epsilon_k H_k(q) e^{-2\beta}(A_k - A_0)e^{2\beta W(q,s^k)}}$$

Now above equation gives,

$$\lambda \sum_{k=1}^{n} \frac{n_k \Delta q}{2\epsilon_k H_k(q) e^{-2\beta}(A_k - A_0)e^{2\beta \quad (q,s^k)}} = 1$$

$$\text{or, } \lambda = \frac{1}{\sum_1^n n_k \Delta q\left[2\epsilon_k H_k(q) e^{-2\beta}(A_k - A_0)e^{2\beta W(q,s^k)}\right]}$$

Now from the above equation, we write,

$$C_k(q) = \frac{n_k/\epsilon_k H_k(q) e^{-2\beta}(A_k - A_0) e^{2\beta W(q,s^k)}}{\sum_{j=1}^{n} n_j\left[\epsilon_j(q) H_j(q) e^{-2\beta}\left(A_j - A_0\right) e^{2\beta \ \ (q,s^j)}\right]}$$

Here we assume that the biased histogram $H_k(q)$ in each umbrella is well estimated and the error function $\epsilon_k(\mathrm{q})$ is equal in all n umbrella windows. Thus,

$$H_k(q) \propto e^{\beta}(A_k - A_0) e^{\beta W(q,s^k) P''(q)}$$

Hence the constant can be rearranged as,

$$C_k(q) = \frac{n_k e^{\beta}\left(A_k e^{-\beta W(q,s^{(k)})}\right)}{\sum_{j=1}^{n} n_j e^{-(q,s^{(j)}}}$$

Therefore the distribution is

$$P''(q) = \frac{\sum_{k=1}^{n} n_k P'_k(q)}{\sum_{k=1}^{n} n_k e^{-\beta(A_k - A_0)} e^{-\beta W(q,s^{(k)})}}$$

$$\text{or, } e^{-\beta(A_k - A_0)} = P''(q) e^{-\beta W(q,s^{(k)})}$$

$$\text{or, } P'' \approx \frac{1}{\sqrt{2\pi\sigma_k^2}} * e^{-(q-q_k)^2}/2\sigma_k^2$$

Where q_k and σ_k^2 are obtaine by MD simulation.

Computation by FORTRAN

Computational simulations are first and time saving technique in modern theoretical physics. Different multi-step calculations are tried to solve by writing computational programming code. Such a basic programming for calculating free energy is FORTRAN program (Raijlich et al., 2001). In the first part of the module, it needs the FORTRAN programming language. To write and run programs in FORTRAN one requires to use a simple text editor to write the code and a compiler to convert it into an executable so that the computer can run. With proper FORTRAN development environment, one can write FORTRAN code, compile it, and run it in a operating system like Linux. Several basic

FORTRAN program (Moses, 1988) like f77, f95 and f90 are available which included as a sample file that calculates the average of the numbers in a file called numbers.txt. Below the code is a line-by-line annotation. Note that in Fortran any line beginning with the character ! is a "comment," which is not translated into the executable program by the compiler. These comments are simply to help the human reader of the code understand it better. All of the code include comments to help or make them more human readable. Linux (Ubuntu) operating system with suitable PC with FORTAN program is used (Yatani et al., 2009) for theoretical estimation of free energy in FORTRAN90 code.

THEORY AND METHODOLOGY

Computational free energy estimation

The conventional method of numerical solution for free energy is non-Boltzmann sampling method (Allen & Tildesley, n.d.) considering the configurationally partition function as,

$$Q_{NVT}^{ew} = \frac{1}{\langle \exp(\beta\Psi) \rangle_{NVT}}$$

As per metropolis Monte Carlo (MC) the potential is negative or small positive. So, umbrella sampling is chosen for estimating the free energy. For any two fluid A and B let us write the potentials $\Psi(\mathrm{r})$ and $\Psi_0(r)$ respectively then the free energy of reference fluid can be written as (Callister, 2007),

$$A - A_0 = -k_B \mathrm{T} \ln\left(\frac{Q}{Q_0}\right) = -k_B \mathrm{T} \ln\left(\langle(\exp(-\beta\Delta\Psi)\rangle_0\right)$$

The free energy depends on different factors and hence can be calculated as, thermodynamically free energy, van-der-Waal free energy, electrostatics free energy, L-J free energy, proper and improper dihedral free energy and solvation free energy. Among them here is the consideration of the dependency of the free energy only with the thermal (β) and density function (conformational) $\rho_0(\Psi)$. Now,

$$\frac{Q}{Q_0} = \int_{-\infty}^{\infty} d(\Delta\Psi) \exp(-\beta\Delta\Psi)\rho_0(\Delta\Psi)$$

Where , $\rho_0(\Delta\Psi)$ is the density representing term of configuration r as,

$\Psi(r) = \Psi_0(r) + \Delta\Psi.$

This ρ_0 contains Boltzmann function $\exp(-\beta\Psi_0)$ Thus the the variation in $\exp(-\beta\Delta\Psi)$ is Boltzmann distribution function (Sevcik, 2017) which favors the configuration with large negative values of$\Delta\Psi$. Torrie and Valleau suggested the general density function as,

$$\rho_w(r) = W(r)\exp(-\beta\Psi_0(r))/\int drW(r)\exp(-\beta\Psi_0(r))$$

Where $W(r) = W(\Delta\Psi(r)$. For this numerical calculation we are considering weighted function for representing the thermal dependency with function $W(-\beta\Delta\Psi)$ which is expressed as,

$$\Delta A = y = W(-\beta\Delta\Psi)$$

Again the density functional dependency is,

$$\Delta A = \rho_0(\Delta\Psi)$$

It is difficult to get weighted function for density dependencyρ_0, even though one can use exponential function as free energy calculation purpose. As it decreases with the rise in temperature, so an exponential function can be used for free energy calculation,

$$\Delta A \propto \exp(\rho_0(\Delta\Psi))$$
$$Or, \Delta A = C\exp(\rho_0(\Delta\Psi))$$

Thus the free energy difference is the product of these two functions as given by equation (Allen & Tildesley, n.d.),

$$\Delta A = -K_bT\ln(\langle\exp(-\beta\Delta\Psi_0)\rangle)$$

FORTRAN code programming

For writing the FORTRAN code of the numerical solution of the free energy let us write the equations,

$y = C\exp(-\beta\Delta\Psi).$

It is also expressed as,

$$y = a0 * \exp\left((\Delta\Psi/(a3 * (x - a1))/a2\right)$$

And for density functional dependency,

$$y = \Delta A = C \exp\left[\rho_0(\Delta\Psi)\right]$$

It is also expressed as,

$$y = a0\exp\left(-(x - a1)^* * 2/a2\right)$$

Here, a0, a1, a2 and a3 are arbitrary constants.

Code I

The code for conformational contribution of free energy is,

!Conformational free energy is expressed as function of temperature.

!$\mathrm{P}(\mathrm{E}_k) = \mathrm{C}\exp\left\{\frac{-(E_k - \langle E_k\rangle)^2}{3NK_B^2T_0^2}\right\}$

!or , $y = a0 * \exp\left(-(x - a1)^2/a2\right)$

!with, $a0 = 1$; Initia lvalue

!$a1 = 315K$; (AverageTemperature),

!$a2 = \sigma^2 * 128$ K

!!programexpo

implicitnone

integer : : i

integer : : x

real : : y

real :: $a0, a1, a2, a3$

open (20, file = "conformational $-FE13.dat$")

$a0 = 10E - 14$

$a1 = 0$

$a2 = 1$

$a3 = 0.00138657$

!change a3 from 0.00118657 to 0.00198657

$do\ x = 250{,}360$

!temperaturerangemayvarieasrequiredhereis jixedwithin 250 to 360

$y = a0 * \exp\left((15/(a3 * (x - a1)))/a2\right)$

write $(*,*)x, y$

write $(20,*)x, y$

!

enddo

endprogram

!graph1isforT315K

!graph 2 for 310 K

!andgraph3 for $320k$

!changeonly05, $1015 - -45$

!changedata

Code II

The code for thermal contribution of free energy is,

!Thermal contribution of Free energy is expressed as function of temperature.

!made by jhulanpowrelcdptu

! $\mathrm{P}(\mathrm{E}_k) = \mathrm{Cexp}\left\{\frac{-(E_k - \langle E_k \rangle)^2}{3NK_B^2T_0^2}\right.$

!or, $\mathrm{y} = \mathrm{a0} * \exp(-(\mathrm{x} - \mathrm{a1})^2/a2)$

! with, $\mathrm{a0} = 1$; Initialvalue

! $a1 = 315K$; (AverageTemperature)

! $a2 = \sigma^{2*}128$ K. !!!programexpo

implicitnone

integer : : i

integer : : x

real :: y

real : : $a0, a1, a2$

open (20, file = "$FE - \mathrm{Therm}_{\ 9}28.dat$")

$a0 = 1$

$a1 = 360$

!changea 1 from 250 to360buthereis fixedvalueat $360a2 = 928$

!changea2from $0.28t0928$

$dox = 0{,}450$

!xisalso fixedat 0,450 onecanvariewithdifferentrangeoftemperarture

$y = a0 * \exp(-(x - a1) ** 2/a2)$

write $(*,*)x, y$

write $(20,*)x, y$

enddo

endprogram

!graph1isfor $250K$

!graph 2 for $255k$

!andgraph3 for 260 ketc

RESULTS AND ANALYSIS

The free energy of a theoretical system with considerable range of parameter is estimated computationally using FORTRAN code. For this purpose we have estimated the conformational as well as thermal free energy separately and then added them to find the resultant free energy of the system.

Plots of conformational and thermal free energy

By considering the parameter of a3 ranging from 0.28 to 928 with interval of 100 the temperature versus free energy graph are plotted for both conformational and thermal contribution of energy.

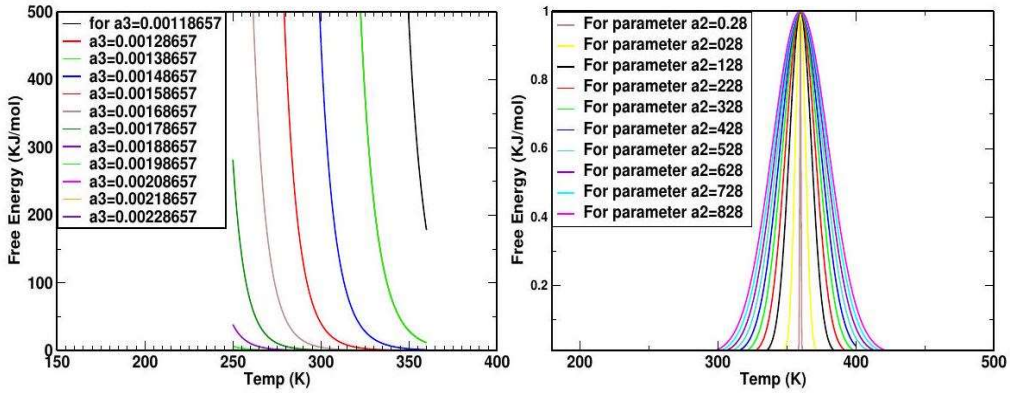


Figure 1: Free energy due to a) conformational and b) due to thermal contribution.

The Figure 1 a) is graph of conformational free energy and figure 1 b) is of thermal contribution of free energy. The different plots are indication of the corresponding free energy. For our convenient we have chosen few of them for estimation of the free energy.

Plots of resultant of free energy

The theoretical plots of conformational and thermal free energy are plotted by Allen and Tildiesely (Figure 2 a)).

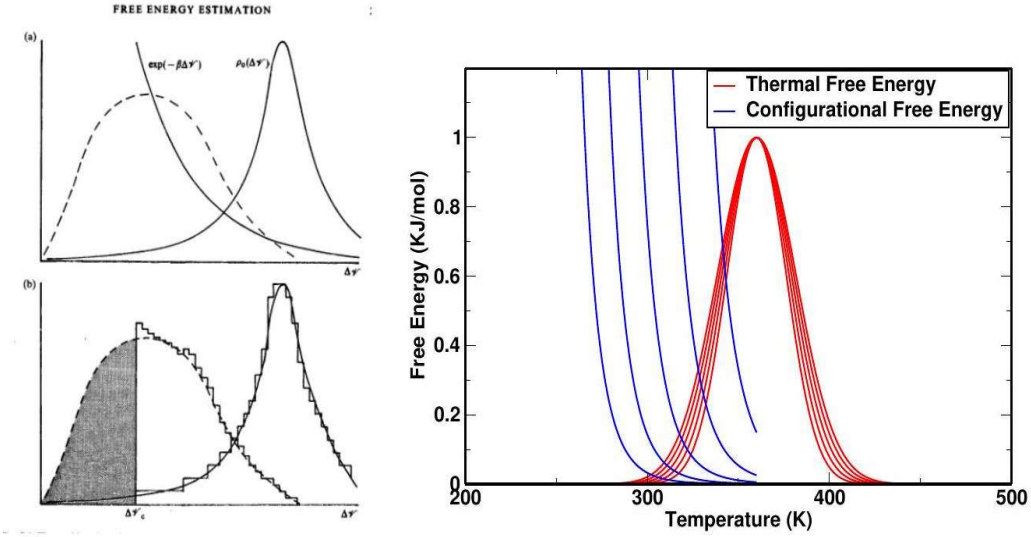


Figure 2: Free energy graph of a) theoretical value b) computational

The resultant plot is shown with dotted line in Figure 2 a) (Allen & Tildesley, n.d.). Selected plots of thermal and conformational free energy are again combines to get the resultant of their free energy as in Figure 2 b).

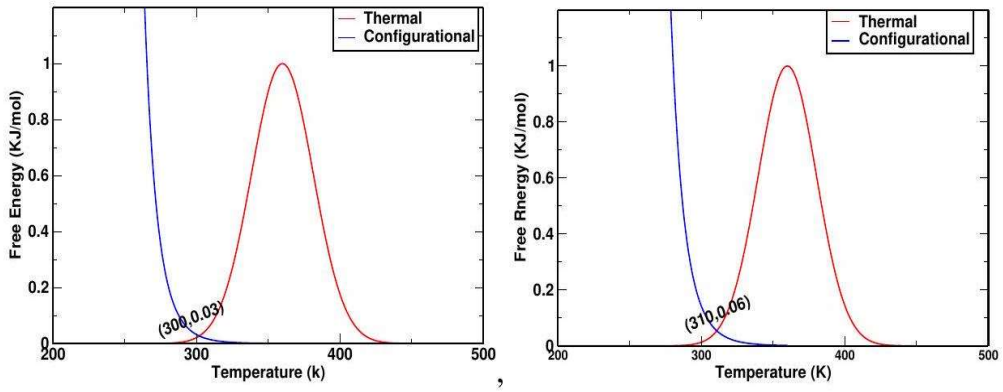
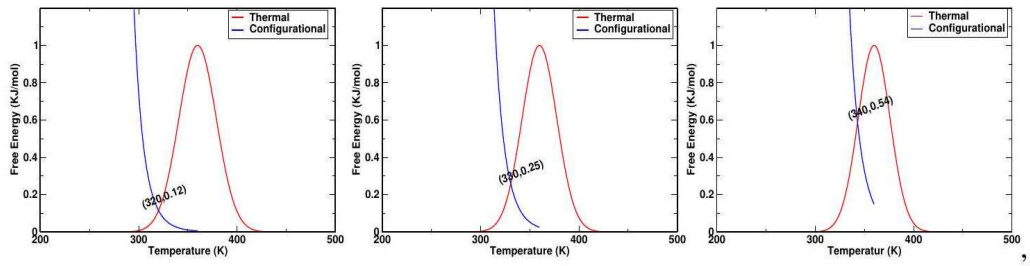


Figure 3: Free energy graph of thermal and conformational contribution at a) 300 K and b) 310 K of temperature.

Individual graphs of thermal and conformational free energy are observed to find their contribution to the total free energy and are tabulated in Table 1. Figure 3a) and 3 b) are the free energy graph giving the energy at temperature of 300 K and 310 K.



a) b) c)

Figure 4: Free energy graph of thermal and conformational contribution at a) 320 K and b) 330 K and at *c*) 340 K of temperatures

Figure 4 a), 4 b) and 4 c) are the individual graph of thermal and conformational contributions of free energy at temperature of 320 K, 330 K and 340 K respectively.

Estimation of Free Energy

Table 1: Estimated values of free energy (KJ/mol) due to conformational and thermal.

From	Temperature (K)	Conf.	Thermal	Total free energy(KJ/mol)
Computation	300	0.03	0.03	0.06
	310	0.06	0.06	0.12
	320	0.12	0.12	0.24
	330	0.25	0.25	0.50
	340	0.54	0.54	1.08
Experimental(n-Butane)	350	–	–	9.03
Theoretical(n-Butane)	350	–	–	-2.75 ± 0.16

The average of the estimated free energy is calculated as $= 0.4\ \mathrm{kJ/mol}$ computationaly. The free energy calculation has also been performed by computational simulation and is agreed with previously performed work (Torrie & Valleau, 1977). The Gibb's free energy for butane has also been estimated from simulation as, $= 2.75 \pm 0.16\ \mathrm{kJ/mol}$ (Torrie & Valleau, 1977). And the free energy of alkane obtained from simulation is, $(2.75 \pm 0.16)\ \mathrm{kJ/mol}$.

CONCLUSION

FORTRAN code can be used for solving the numerical problem like free energy estimation in molecular dynamics. From graph the average of the computational free energy of a theoretical system is estimated as $0.4\mathrm{KJ/mol}$, which is closed and within the range of theoreticaly calculated (Table 1) and experientally (Table 1) observed free energy.

DATA AVAILABILITY STATEMENT

All data that support the findings of this study are included within the article (and any supplementary files).

AUTHORS CONTRIBUTION

JP carried out all the works and idea of study. Authors read and approved the final manuscript. All data and paper are available with the correspondent author.

REFERENCES

Allen, M., & Tildesley, D. (1987). *J. computer simulations of liquids 1987*. Oxford University Press, New York.

Callister, W. D. (2007*). Materials science and engineering: An introduction. John Wiley & Sons.* Inc.: Utah.

DeVries, P. L., & Hamill, P. (1995*). A first course in computational physics. American Association of Physics Teachers*. USA.

Jarzynski, C. (1997, Apr). *Nonequilibrium equality for free energy differences. Phys. Rev. Lett., 78, 2690-2693.* Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett. 78.2690 doi: 10.1103 /PhysRevLett. 78.2690.

Kittel, C., & McEuen, P. (1976). *Introduction to solid state physics, vol 8 wiley.* New york.

Moses, G. A. (1988*). Engineering applications software development using fortran 77. John Wiley & Sons, Inc.*

Raijlich, V., Wilde, N., Buckellew, M., & Page, H. (2001*).* Software cultures and evolution*. Computer*, 34(9), 24-28.

Sevcik, C. (2017*).* Caveat on the Boltzmann distribution function use in biology. *Progress in Biophysics and Molecular Biology*, 127, 33-42.

Stradner, A., Foffi, G., Dorsaz, N., Thurston, G., & Schurtenberger, P. (2007*).* New insight into cataract formation: Enhanced stability through mutual attraction. *Physical Review Letters*, 99(19), 198103.

Torrie, G. M., & Valleau, J. P. (1977). Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics,* 23(2), 187 − 199.

Yang, M., & MacKerell Jr, A. D. (2015*).* Conformational sampling of oligosaccharides using Hamiltonian replica exchange with two-dimensional dihedral biasing potentials and the weighted histogram analysis method (wham). *Journal of Chemical Theory and Computation,* $11(2), 788 - 799$.

Yatani, K., Chung, E., Jensen, C., & Truong, K. N. (2009). Understanding how and why open source contributors use diagrams in the development of Ubuntu. *In Proceedings of the Sigchi Conference on Human Factors in Computing Systems* (995-1004).