

University of Alberta

Atomistic Simulations for Investigating Structural Stability
and Selecting Initial Adsorption Orientation of Lysozyme and Apo- α
-Lactalbumin at Hydrophobic and Hydrophilic Surfaces

by

Siriporn Pansri

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

in

Materials Engineering

Department of Chemical and Materials Engineering

©Siriporn Pansri

Fall 2012

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Abstract

Molecular dynamics (MD) simulations were performed to investigate the structural stability of lysozyme and apo- α -lactalbumin under physiological pH and solution. Upon introduction to the solution, apo- α -lactalbumin showed lower stability than lysozyme with higher backbone root mean square fluctuation (RMSF) of the exposed residues and C-terminus, including secondary structure transition from α -helices to turns in residues 105-110. However, no noticeable changes were observed in the secondary structure of lysozyme during the simulations. Subsequently, molecular mechanics (MM) simulations were carried out to determine the preferred orientation for adsorption of these proteins at poly (ethylene oxide) (PEO) films capped with hydroxyl and methoxy end-groups, based on the Lenard-Jones (L-J) potential. Both proteins preferred to initially adsorb at hydrophilic surface with its side-on orientation which clef faces sideways. At hydrophobic surface, the preferred orientation of lysozyme became a back-on orientation which clef faces outward, whereas that of apo- α -lactalbumin remained a side-on orientation.

Acknowledgements

First of all, I would like to thank my supervisor, Dr. Larry Unsworth, and my co-supervisor, Dr. Andriy Kovalenko, for giving me an opportunity to do a project in the area of biophysics. It has been very fun.

Secondly, I would like to take this opportunity to thank to Dr. Nikolay Blinov for all of his supports, ranging from carrying my computer, correcting and helping about all scripts, and giving advice with unlimited patience. This work would not be possible to accomplish without all of his supports.

Thirdly, I would like to thank my co-supervisor again for all computing resources including personal computing and access to computing resources at Westgrid and the NINT HPC cluster. Without these resource, none of the lengthy simulations would have succeeded.

More importantly, I would like to thank the Royal Thai Government for all financial supports under the program Strategic Scholarship for Frontier Research Network (SFR) of Thailand's Commission on Higher Education. The scholarships financial supports, including tuition fees, cost of living, health plan, and air fare, made it possible for me to have different and valuable experiences in Canada.

Last but not least, I would like to thank my parents, my brothers and all my relatives for all the long distance support. Also, I would like to thank all of my Thai friends from the Thai Student Association at the University of Alberta for their encouragement.

Contents

Title Page

Abstract

Acknowledgements

Contents

List of Tables

List of Figures

List of Abbreviations

1 Introduction	1
1.1 Protein Adsorption	1
1.2 Main Driving Forces for Protein Adsorption	1
1.3 PEO and Its Ability to Resist Protein Adsorption	5
1.4 Lysozyme and α -lactalbumin	7
1.5 Atomistic Simulations	8
1.6 Scope of the Thesis	8
2 Backgrounds	10
2.1 Adsorption Thermodynamics	10
2.2 Molecular Modelling	11
2.2.1 Overview	11
2.2.2 Basic Concepts of Molecular Dynamics simulations .	12
2.2.3 Molecular Mechanics Concepts	23
2.3 Summary	25
3 Molecular Dynamic Simulations of Human Lysozyme and	

Apo- α-Lactalbumin	26
3.1 Introduction	26
3.2 Preliminary studies of X-ray crystal structure of Human lysozyme and α -Lactalbumin	27
3.3 Structural Analysis of Human Lysozyme and Human Apo- α -Lactalbumin Using Molecular Dynamics Simulations . . .	35
3.3.1 Methodology	35
3.3.2 Results and Discussion	38
3.4 Summary	49
4 Molecular Mechanics Simulations of Lysozyme and Apo-α-Lactalbumin Adsorption at Hydrophobic and Hydrophilic Surfaces	50
4.1 Introduction	50
4.2 Methodology	52
4.2.1 Description of Polymer, Protein Models and Simulations	52
4.2.2 Potential Energy	60
4.3 Results and Discussion	61
4.3.1 Effect of a Protein Orientation on Adsorption of Lysozyme at Hydrophobic PEO-OCH ₃ surface . . .	61
4.3.2 Effect of a Protein Orientation on Adsorption of Lysozyme at Hydrophilic PEO-OH Surfaces	68
4.3.3 Effect of Orientation on Adsorption of Apo- α -lactalbumin at Hydrophobic PEO-OCH ₃ Surface	73
4.3.4 Effect of Orientation on Adsorption of Apo- α -lactalbumin at hydrophilic PEO-OH Surface	77
4.4 Summary	82
5 Summary	83
5.1 Current Thesis	83
5.2 Future work	85

Appendix A

95

text page

List of Tables

2.1	Method of steepest decent ⁶⁸	24
2.2	Method of conjugate gradient ⁶⁹	25
3.1	Secondary structure of human lysozyme and human holo- α -lactalbumin assigned using the VMD software.	30
4.1	Different orientations of lysozyme (left) and apo- α -lactalbumin (right). $R_3 90^\circ$, $R_3 180^\circ$, $R_3 270^\circ$, $R_3 360^\circ$ stand for 90° , 180° , 270° and 360° clockwise rotation about the 3rd axis, respectively. $R_1 90$ and $R_1 270$ stand for 90° and 270° clockwise rotation about the 1st axis, respectively. In cartoon representations, purple, blue, cyan, white, yellow and gold colours present α helix, helix 3/10, turn, coil, main $-\beta$ sheet and partial- β sheet conformations, respectively.	56
A.1	Atom name, type and atomic partial charge assigned by MOPAC (bcc) for PEO-OCH ₃	97
A.2	Atom name, type and atomic partial charge assigned by MOPAC (bcc) for PEO-OH	98
text	page

List of Figures

2.1	Notations of angles and planes accounted for an empirical force field. ¹³	14
2.2	Molecular dynamics algorithm	17
2.3	Periodic boundary condition(PBC) ¹⁴	21
2.4	Oscillation or zig-zag patterns of move in a long narrow valley shown as contour lines. ²	24
3.1	a) Shape of lysozyme. Solvent accessible surfaces corresponding to hydrophilic and hydrophobic residues are shown in red and blue, respectively. b) Secondary structures of lysozyme; 4 α helices (α_1 , α_2 , α_3 , α_4), 4 helices 3/10, β sheets (β_1 , β_2 and β_3), turns and coils depicted as purple and blue and yellow, cyan and white colours, respectively, cystein residues illustrated as yellow van der Waals (vdW) spheres. Residues marked by arrows are highly flexible residues. (indicated as white arrows in Figure 3.3 a). .	28

3.2	a) Shape of α -lactalbumin. Solvent accessible surfaces corresponding to hydrophilic and hydrophobic residues are shown in red and blue colours, respectively. b) Secondary structures of holo- α -lactalbumin; 4 α helices (α_1 , α_2 , α_3 , α_4), 5 helices $3/10$, β sheets (β_1 , β_2 and β_3), turns and coils depicted as purple, blue, yellow, cyan and white colours respectively, cysteine residues illustrated as yellow van der Waals (vdW) spheres. Residues marked by arrows are highly flexible residues. (indicated as white arrows in Figure 3.3 b).	29
3.3	Atomic fluctuation, secondary structure of each residue and disulfide bridges formed in lysozyme a) and holo- α -lactalbumin b) labeled as arrows linking cysteine residues. Different values of temperature β -factor are illustrated by different intensity of colour, with darker colour indicating higher values, and consequently higher flexibility. Domains with noticeably high fluctuations marked with white arrows. In secondary structure maps, α helices, helix 3_{10} , main β sheets, partial- β sheets, turns and coils are depicted with purple, pink, yellow, gold, green and white, respectively.	31
3.4	Charts representing values of solvent accessible surfaces of lysozyme a) ³³ and holo- α -lactalbumin b) ³² , different scales of colour indicating different percentage of solvent accessible surfaces: white colour indicating 100 % solvent accessible surfaces, while black colour indicating totally buried surfaces.	34
3.5	Time required to carry out 1 ns MD simulations for human lysozyme as a function of number of CPUs used. 32 is optimum number of CPUs for this system.	38
3.6	a) Backbone RMSD with respect to crystal structure of each protein, b) Backbone RMSD with respect to structure of each protein after minimization runs.	41

3.7	Backbone RMSF with error bars of lysozyme and apo- α -lactalbumin. Regions with highly fluctuating residues are depicted as green colour in cartoon representations of lysozyme (top right) and apo- α -lactalbumin (top left).	43
3.8	Time evolution of secondary structure of lysozyme, purple, pink, green, yellow, gold and white represent α helices, helices 3/10, turns, main- β sheets, partial- β sheets and coils, respectively.	45
3.9	Time evolution of secondary structure of apo- α -lactalbumin, purple, pink, green, yellow, gold and white represent α helices, helices 3/10, turns, main- β sheets, partial - β sheets and coils, respectively.	46
3.10	Radius of gyration of lysozyme and apo- α -lactalbumin.	48
4.1	a) A complex system composed of a protein translated in z direction with an increment of 0.5 Å and HS-PEO-OCH ₃ translated in x-y plane depicted in the figure (there are three different things here: (i) protein translation as a part of MM protocol (ii) polymers translation as a part of the procedure used to build a surface (iii) chain length of OCH ₃ group \sim 2 Å attached to O10 atom). CH ₃ groups, sulfur atoms, O10 atoms and PEO chains labeled as cyan, yellow, red vdW spheres and cyan lines, respectively. b) Graphical explanation of different parts of lysozyme molecule c) Graphical representation of different parts of apo- α -lactalbumin molecule.	54
4.2	L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OCH ₃ surface.	62
4.3	Different initial adsorption orientation of lysozyme at PEO-OCH ₃ surface. Red, blue and white colours represent acidic, basic and other residues, respectively.	64

4.4	a) Residue distribution of each lysozyme orientation over PEO-OCH ₃ surface including indications of the most interacting region (14 Å) and approximate L-J cut off region (18 Å) as well as ccm of CH ₃ groups (2 Å). b) Summarization of total number of residues lying in the most interacting region.	67
4.5	L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OH surface.	68
4.6	Different initial adsorption orientation of lysozyme at PEO-OH surface. Red, blue and white colours represent acidic, basic and other residues, respectively.	71
4.7	a) Residue distribution of each lysozyme orientation over PEO-OH surface with indications of approximate L-J interacting region and the most interacting region at the distance of 16 and 12 Å, respectively. b) Summarization of residues lying in most interacting region.	72
4.8	L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OCH ₃ surface.	73
4.9	Different initial adsorption orientation of apo-α-lactalbumin at PEO-OCH ₃ surface. Red, blue and white colours represent acidic, basic and other residues, respectively.	74
4.10	a) Residue distribution of each apo-α-lactalbumin orientation over PEO-OCH ₃ surface including indications of the most interacting region (14 Å) and approximate L-J cut off region (18 Å) as well as ccm of CH ₃ groups (2 Å). b) Summarization of total number of residues lying in the most interacting region.	75
4.11	L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OH surface	78
4.12	Different initial adsorption orientation of apo-α-lactalbumin at PEO-OH surface. Red, blue and white colours represent acidic, basic and other residues, respectively.	80

4.13 a) Residue distribution of each apo- α -lactalbumin orientation over PEO-OH surface with indication of approximate L-J and most interacting regions at the distance of 16 Å and 12 Å , respectively. b) Summarization of residues lying in most interacting region. 81

text page

List of Abbreviations

ns	nano second
ps	pico second
fs	femto second
g/cm ³	gram per cubic centimeter
LJ	Lennard Jones
MD	Molecular dynamics
MC	Monte Carlo
MM	Molecular Mechanics
PEO	Poly(ethylene oxide)
PEG	Ply(ethylene glycol)
PBC	Periodic Boundary Conditions
PS	Polysterene
vdW	van der Waals
EL	Electrostatic
kcal/mol	kilo calorie per mol
NTP	Isothermal-isobaric ensemble
NTV	Canonical ensemble
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
EL	Electrostatic
QM	Quantum mechanics
SMD	Steered molecular dynamics
VMD	Visualize molecular dynamics
Ala	Alanine
Val	Valine
Leu	Leucine
Ile	Isoleucine
Phe	Phenylalanine
Trp	Tryptophan
Met	Methionine
Pro	Proline

Gly	Glycine
Ser	Serine
Thr	Threonine
Cys	Cysteine
Tyr	Tyrosine
Asn	Asparagine
Gln	Glutamine
Asp	Aspartic acid
Glu	Glutamine acid
Lys	Lysine
Arg	Arginine
His	Histidine
SAM	Self-assembled monolayer

Chapter 1

Introduction

1.1 Protein Adsorption

Generally, once biomaterials are in contact with biofluids, such as whole blood or plasma, proteins adsorb within the first few minutes and form a layer. Moreover, multilayers may form as proteins may adsorb upon the previously adsorbed layer. Non-specific protein adsorption is crucial to the therapeutic efficacy of the blood contacting material, as it is well known that proteins at surfaces can lead to the initiation of several host responses. Cells may function with the biomaterials through active sites provided by this layer of adsorbed proteins. Moreover, these adsorbed layers may also provide conditions suitable for bacterial colonization, leading to infection and the eventual failure of the biomaterials ability to function in the desired therapeutic capacity it was designed for.

1.2 Main Driving Forces for Protein Adsorption

Typically, adsorption of proteins on a selected surface in an aqueous solution is influenced by several factors related to the combined properties of the surface, proteins, and solution.¹ For example, it has been theoretically reported that lysozyme tends to adsorb on hydrophobic CH_3 SAM surfaces rather than on hydrophilic oligo (ethylene glycol) surfaces.³ Moreover, it

has also been reported that denatured protein, particularly α -lactalbumin without bound Ca^{+2} ions, adsorbed to a greater extent on stainless steel coated with polyethylenimine (PEI) than it did under its native state with bound Ca^{+2} ions.⁵ The protein without bound Ca^{2+} ions adsorbed less on PEG-coated surfaces of various molecular weights.³ In addition, it has also been reported that Arg^+ at pH 7.0 was adsorbed on a SAM-COOH surface, whereas the residue was repelled from the SAM-NH⁺³ surface.³ Furthermore, it has also been reported that at high salt concentrations, in an aqueous solution of pH 7.0, the extent of adsorbed lysozyme at negative polystyrene decreased compared with the adsorbed extent of the protein at a low salt concentration.⁸

Several factors affect protein adsorption on a given surface. As a result, the mechanisms associated with protein adsorption are complex.¹ However, the investigation of these mechanisms is limited on an experimental scale. Therefore, theoretical studies have been conducted. Theoretically the main driving forces for protein adsorption processes may be classified as Van der Waals (vdW), electrostatic (EL) and hydrophobic interactions, hydrogen bonding, and subprocesses directly related to a proteins stability.

vdW interaction plays a role in all adsorption processes^{6,34} with an effective range varying from a few to $\sim 20 \text{ \AA}$.²⁷ It is also known that vdW is dominated only where hydrophobic and EL interactions are weak, such as in vacuum or in continuum medium.²⁰ vdW interactions can be described by various types depending on the potential well selected, for instance, there are two contributions of vdW interaction based on the Lennard-Jones (L-J) potential well namely steric and dispersion terms. The dispersion force being analogous to an instantaneous dipole-induced dipole interaction between two atoms. The strength of the interaction being dependant on the size of the atom, since the larger atom poses a higher polarizability and more dispersed electron clouds. On the other hand, the steric term arising from the Pauli exclusion principle observed when two atoms are repelled from each other when they are closer than the sum of their two radii, as a result of overlapping of electron clouds.

EL interaction plays a major role in protein adsorption especially when charges of a protein and a surface are non-zero. According to Coulombs Law a charged particle generally attracts an oppositely charged particle, while repelling a likely charged particle. Consequently, if a protein and bio-material surface have opposite charges, EL interaction can also promote adsorption with effective range of ~ 60 nm.²⁷ For instance, Norde et al.⁸ reported that lysozyme, which is positively charged at pH 7.0 was strongly adsorbed at negatively charged poly (styrenesulphonate). Another example reported by Norzita et al.⁵ was that a positively charged lysozyme was unfavorable and a negatively charged α -lactalbumin was favorable to adsorb at positively charged polyethylenimine (PEI) (pH=7.2).

Hydrophobic interactions can be long-range with an effect over ~ 10 nm.²⁷ The interaction is well known as being the impetus for the aggregation of hydrophobic solute in an aqueous solution, which plays an important role in several processes such as the self-assembly process, protein stabilization, and protein adsorption in an aqueous solution.²⁷ Generally, a hydrophobic solute does not form hydrogen bonds with water molecules; therefore, its hydration shell is not as strong as that around a hydrophilic solute, where hydrogen bonds form between the solute and water molecules. Without the formation of solute-water hydrogen bonds, the hydration shell of the hydrophobic solute is easily perturbed during protein adsorption. The destruction of the hydration shell gives rise to the amount of the free energy of adsorption in terms of favorable entropy due to the reordering of water molecules. Moreover, several authors,^{15,19} also reported that density fluctuation around the hydrophobic solute was very high. Since, fluctuation and compressibility are theoretically directly related to each other¹², the area around the hydrophobic solute was highly compressible. High density fluctuation surrounding a hydrophobic solute has been used to explain why adsorption of proteins at hydrophobic surfaces induced more conformational changes of the adsorbed proteins than that at hydrophilic surfaces.^{28,29} Since high density fluctuation above hydrophobic surfaces induced motion of the adsorbed protein over the surface and consequently large conformational change of the adsorbed pro-

teins.^{10,11,20,22}

Hydrogen bonding is a short range interaction with an effect over $\sim 1\text{-}2 \text{ \AA}$.²⁷ As a result, number of hydrogen bonds formed between water molecules and a solute affects strength of a hydration shell surrounding the solute. The higher the number of hydrogen bonds formed between the solute and water molecules, the stronger the hydration shell is. Strong hydration shell surrounding a solute cannot be perturbed easily while another solute approaches.^{30,47} To consider adsorption processes, if the biomaterial surface forms a large number of hydrogen bonds with water molecules, a hydration shell surrounding the biomaterial surface will be very strong. Hence, when a protein approaches to the biomaterial surface, the strong hydration shell will generate repulsive force at the protein.^{30,47} Moreover, strong hydration shell can also reduce favorable entropy resulted from perturbation of the hydration shell while the approaching protein may induce restructuring of water molecules in the shell. Therefore, this can reduce adsorption of a protein at biomaterial surface.²⁷ Accordingly, hydrogen bonding is considered one of the main driving forces of protein adsorption.²⁷

Protein stability is also an important factor that can generate favourable entropy during the adsorption process. Soft proteins, such as α -lactalbumin, that have a low structural stability, have a strong tendency to change their conformation upon adsorbing compared to hard proteins such as lysozyme; thus, the adsorption of the soft protein, especially on hydrophobic surfaces, usually increases the system entropy through the release of water molecules associated with both the protein and the biomaterial.¹ Consequently, this gives rise to the extent of adsorption energy. Moreover, it has also reported that hydrophobic residues of denatured proteins are exposed to an aqueous solution more than when the protein is in its native structure.¹⁵ Accordingly, the denatured protein can also promote further adsorption *via* subsequent hydrophobic interactions.

1.3 PEO and Its Ability to Resist Protein Adsorption

Polyethylene oxide (PEO, $\text{H}-(\text{OCH}_2\text{CH}_2)_n-\text{OH}$) is an electrically neutral polymer that is highly soluble and very mobile in aqueous solutions at low temperatures, compared to other polymers.²¹ As a result of these interesting properties, PEO has been selected as one of the best candidates for non-fouling surfaces. The hydrophilic nature of PEO has been of interest to several authors.^{41,42,43} The unique hydrophilicity of PEO was explained by Ronald et al.⁹ several years ago to be a result of a good fit between the ether oxygen of PEO and the tetrahedral lattice of water. The ether oxygen of PEO being able to fit into the water lattice without significantly disturbing the tetrahedral network of water (oxygen-oxygen distance = 2.85 Å between neighbors and 4.7 between next-nearest neighbors); hence, Ronald et al.⁹ concluded that PEO was highly soluble. PEOs high degree of hydrophilicity also corresponds to a large number of hydrogen bonds formed in its hydration layer. The electrostatic attraction between the negative potential of the lone pair of electrons of ether oxygen and the positive potential of the OH bond of water was believed to be the driving force behind the formation of hydrogen bonds, as reported by Weitz et al.⁴² Moreover, it was reported that the hydration shell of PEO was very stable and well ordered with the contribution of negative enthalpy and entropy of hydrogen bonds formation in the hydration shell at room temperature or even higher temperatures (450 K).^{41,43} The degree of hydrogen bond formation also depends on the density of the PEO chains. If the chains are too dense, the number of hydrogen bonds determined in the hydration shell is generally fewer than in cases of optimum chain density.^{41,43} In contrast, in cases of low PEO chain density, even water can penetrate well into the PEO chains. The number of hydrogen bonds formed is low because of limitation of interacting sites for hydrogen bonds formation.³

Another interesting property of PEO is its mobility. PEO is known to be a very mobile polymer.²¹ This may be explained by its structure having no bulky chemical groups, being comprised of single bonds along

the chemical backbone, and by its electric neutrality; thus, this leads to a small effect of steric and electrostatic hindrance. Moreover, Davies et al.¹⁶ reported from dipole relaxation that reorientation in PEO chain is a marked freedom and the carbon-oxygen linkage of the PEO chain behaves like a ball-and-socket joint in which carbon or cation atom behaves like a ball which is allowed to move freely, whereas oxygen behaves like a socket to accommodate the carbon atom. They also found that the average movement of longer units were more freely to move than a single unit.¹⁶ The flexibility of PEO has attracted several authors. For example, Zeng et al.³ studied the flexibility of ethylene oxide (OEG), and reported that OEG SAM was more flexible and disordered than SAM terminated with OH and $\text{CH}_3(\text{S}(\text{CH}_2)_4(\text{EG})_4)\text{OH} > \text{S}(\text{CH}_2)_{11}\text{OH} > \text{S}(\text{CH}_2)_{11}\text{CH}_3$). The disordered structure of OEG SAM enhanced water penetration into the OEG film to form hydrogen bonds with ether oxygens of the OEG SAM; thus, in the study the number of hydrogen bonds formed in the hydration layer of OEG SAM was more than those formed in case of the other two SAMs. Moreover, it is reported that as the PEO chain length increased, chain flexibility increased, as experimentally reported by Han et al.⁴⁸

PEOs ability to resist protein adsorption is known to be the result of its two unique properties, described above. For example, Jie et al.³ observed lysozyme adsorption on $\text{HS}(\text{CH}_2)_4(\text{EG})_4\text{OH}$, $\text{HS}(\text{CH}_2)_{11}\text{OH}$ and $\text{HS}(\text{CH}_2)_{11}\text{CH}_3$ surfaces. They found that total forces exerted on the lysozyme were mainly generated from the water shell, and that the magnitude of the forces was closely related to the number of hydrogen bonds in the PEO hydration shell. PEOs packing density also affected its ability to resist protein adsorption in terms of the stability of the hydration shell, as mentioned above. Furthermore, Jeon et al.⁴⁹ also reported that the best condition for preparing a PEO non-fouling surface was at an optimum PEO chain density with a longer chain length. However, Unsworth et al.¹⁷ reported that the ability to resist adsorption of fibrinogen was not different when the PEO chain length increased (MW of 750 and 2000) at the optimum chain density (0.5 nm^{-2}). Moreover, Norzita et al.⁵ reported that negatively charged apo- α -lactalbumin adsorbed to surfaces modified with

PEI-PEO (350, 2000, and 5000 MW) comb-like structures PEO surfaces, and silica coated on bare stainless steel. These studies indicate that even PEO, which is very highly hydrated and mobile cannot prevent adsorption of all types of proteins.

1.4 Lysozyme and α -lactalbumin

Lysozyme and α -lactalbumin have been studied since the middle of the twentieth century and, thus, are considered well-characterized proteins. It is interesting that these two proteins share a similar size, shape, and folding patterns, as well as number of disulfide bonds. Lysozyme and α -lactalbumin have an approximately ellipsoid shape with nearly the same molecular weight, 14 kDa. In addition to the similarities in size and shape, their three dimensional folds are also homologous, with a cleft separating the larger helical portion and the smaller portion of β -sheet. The similarity in folding patterns is believed to be a result of 35-40 % in sequence similarity.³⁵ Lastly, both proteins are also composed of four disulfide bonds. Despite the homological characteristics of both proteins, their functions are far different. Lysozyme is involved with binding and cleaving the glycosidic bond in sugar, while α -lactalbumin takes part in lactose synthesis.³⁵ Moreover, lysozyme and α -lactalbumin also have different stabilities. The structure of lysozyme was reported to be more stable than that of α -lactalbumin with $\Delta G_{Thermal}$ and $\Delta G_{Denaturant}$ of +4.1 and +4.0 in case of lysozyme and +1.5 ,and +1.9 in case of α -lactalbumin²⁸, where $\Delta G_{Thermal}$ and $\Delta G_{Denaturant}$ are changes in free energy as a result of denatured and native states of a protein due to temperature and pH, respectively.²⁸ In addition, it is reported that α -lactalbumin is a bound Ca^{2+} protein at primary and secondary sites.³⁵ Besides, only the bound Ca^{2+} ion at the primary site plays an important role in the structural stability of α -lactalbumin. (More information can be found in Chapter 3.). Additionally, the bound protein with Ca^{2+} ion at this primary site shows higher stability than that without a bound Ca^{2+} ion.³⁵ It is worth noting that α -lactalbumin under the unbound Ca^{2+} ions is called apo- α -lactalbumin,

while the protein under its native conformations (with bound Ca^{2+} ions) is called holo- α -lactalbumin. Moreover, their isoelectric points (pI) are also different: lysozyme being around 11, while holo α -lactalbumins is ~ 4.5 .

1.5 Atomistic Simulations

Classical based atomistic modeling is generally used to examine behaviours of big systems such as a system of biomolecules containing more than 10^4 atoms. Molecular mechanics (MM), Monte Carlo (MC), and molecular dynamics (MD) are three well-known types of classical based simulations. Among the three potential-based methods, MM is different from the other two methods in that during simulations, the effect of thermal motion is not taken into account. Furthermore, MM is also called the energy minimization or geometry optimization method because it is commonly used to minimize the potential energy of the system. Additionally, MM has been used in place of MC simulations to select the most preferable- interacting sites between biomolecules like antigen-antibody, enzyme-substrate and protein-surface because of computationally cheaper cost.^{51,52,53,54,55} Due to the fact that MC and MD techniques incorporate the effect of temperature, they can be used to calculate thermodynamic properties. All atoms in MD simulations are moved following Newton's second law of motion to reproduce the time evolution of the system in the phase space. Unlike MD, the atoms in MC are moved following a a probalilistic algorithm to recapitulate the statistics of energy and spatial distributions for a canonical ensemble in thermal equilibrium. Therefore, to deal with dynamical properties of a system, MD is required rather than MC.^{13,14}

1.6 Scope of the Thesis

Since the internal stability of proteins can also play a role in adsorption process. Soft proteins tend to be denatured more than hard proteins during adsorption, especially on hydrophobic surfaces. Moreover, both soft and

hard proteins also prefer to adsorb on hydrophobic surfaces rather than hydrophilic ones. However, mechanisms of these still cannot be explained clearly in experiments. The present thesis is aimed at elucidating these questions by investigating the adsorption mechanism of two model proteins, namely lysozyme and α -lactalbumin on hydrophobic and hydrophilic surfaces by using atomistic simulations.

Chapter 2

Backgrounds

2.1 Adsorption Thermodynamics

Generally, when a protein in aqueous solution exposes to a solid surface, a spontaneous adsorption of the protein at the solid surface is observed. In theory, protein adsorption at a solid surface can occur spontaneously at constant temperature and pressure if the change in Gibbs free energy (ΔG) on the adsorption is less than zero :⁴

$$\Delta G = \Delta H - T \Delta S < 0 \quad (2.1)$$

where, H is the enthalpy, S is the entropy and T is the temperature. Behaviours of an adsorbed protein at the solid surface are governed by net interactions between the adsorbed protein, the solid surface and the aqueous solution.¹ These interactions may arise from non-bounded interactions such as electrostatic (EL), van der Waals (vdW) and hydrophobic interactions or bounded interactions like hydrogen and covalent bonding which are typically accounted for in the enthalpy term. Additionally, other effective interactions in dense medium such as secondary packing of the adsorbed protein, cavity formation and degree of solvation are also accounted for in both the enthalpic and entropic terms.

2.2 Molecular Modelling

2.2.1 Overview

Molecular simulation methods can generally be divided into three groups, namely, molecular mechanics (MM), Monte Carlo (MC) and molecular dynamics (MD) simulations. Among the three molecular simulation methods, MM is different from the other two methods in that during simulations thermal motion is not taken into account, while it is in the other two methods. Furthermore, MM is also called energy minimization or geometry optimization method because it is commonly used to minimize potential energy of the system. Additionally, MM is usually applied to minimize initial structure of a system before performing MD simulations. Moreover, recently MM simulations have also been carried out to predict the most preferable adsorption orientation of a protein at a given surface based on global energy minimum instead of MC simulations since they are computationally cheaper.^{51,52,53,54,55} Due to MC and MD take effects of temperature into account during simulations; hence, they are used for calculations of thermodynamic properties. In MC, atoms are moved according to various time independent probabilistic algorithm to recapitulate the canonical ensemble statistics in thermal equilibrium, whereas in MD the atoms are moved following the Newton's second law of motion to reproduce the time evolution of the system in the phase space. As a result, transport properties such as rate of an interaction, and diffusion coefficient can be determined only by performing MD simulations. This is one of the reasons why MD simulation is commonly used to investigate the behavior of bimolecular systems. In view of the above, both MM and MD simulations will be carried out in this thesis. The two methods are briefly described in this chapter, whereas further information on molecular modeling can be found elsewhere.^{13,14}

2.2.2 Basic concepts of Molecular Dynamics simulations

In MD simulations, each particle moves according to the Newtons second law of motion, Equation (2.2), where the force is related to the potential energy by Equation (2.3).

$$\vec{F}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \quad (2.2)$$

$$\vec{F}_i(\vec{r}_i) = -\vec{\nabla} U_i(\vec{r}_i) \quad (2.3)$$

According to Equation (2.2) and (2.3), \vec{F}_i stands for all forces acting on the atom i with constant mass m_i , \vec{r}_i is a position vector of the atom i and $U(\vec{r}_i)$ is potential energy of the atom i .

It is often assumed that the potential energy is mostly defined by pairwise interactions between particles. The interactions which involve a larger number of particles (such as three-particle or higher order interactions) are usually weaker than pairwise interactions, they are often included effectively rather than explicitly in empirical force fields. Particular interactions contributing to the potential energy will be described in details in the next subsection.

Empirical force field

The forces acting on atoms in the course of MD simulations arise from both bonded and non-bonded interactions. These include contributions due to bond-stretching, bond-bending, dihedral and improper dihedral angles, as well as van der Waals (vdW) and Coulomb interactions. An equation that includes all parameters used for calculations of potential energy of a specified system is often called the equation of empirical force field. The parameters included in this equation are calibrated by fitting to the experiment and/or to the results of the QM calculations. Presently, there are many types of force fields that exist for molecular modeling, and include but are not limited to CHARMM, AMBER, and GROMOSS force fields. The same force field may not be appropriate even for the same biomacromolecule system within different environments (such as solvent

composition or thermodynamic state), and a proper force field should be chosen carefully for a particular situation.

Typically, a standard form of a modern force field equation is given by Equation (2.4).

$$\begin{aligned}
 U_{total} = & \sum_{bond} U_{bond} + \sum_{angle} U_{angle} + \sum_{dihedral} U_{dihedral} + \sum_{improper} U_{imprpper} \\
 & + \sum_{vdW} U_{vdW} + \sum_{Coulomb} U_{Coulomb}
 \end{aligned}
 \tag{2.4}$$

In Equation (2.4), the first four terms describe bonded interactions, while the last two terms represent van der Waals and Coulomb (non-bonded) interactions. These terms are briefly described below.^{13,27}

Bond stretching (the first term in Equation (2.4)) can be described by the Hooks law applied for each two particles connected by the covalent bond (see Figure 2.1). Mathematically, this interaction is described by Equation (2.5), where k is a spring constant, and b and b_0 are instantaneous and equilibrium bond lengths, respectively.

$$U_{bonds} = \frac{1}{2}k_b(b - b_0)^2
 \tag{2.5}$$

Bond Bending contribution to the potential energy can also be written in the harmonic form. This term describes variations in the potential energy due to change of angle (θ) between three consecutives covalently bound atoms (such as atoms 1, 2 and 3 in Figure 2.1). Mathematical expression for this type of interactions is given by Equation (2.6), where θ_0 and θ are equilibrium and instantaneous angles between three atoms, respectively.

$$U_{angle} = \frac{1}{2}k_\theta(\theta - \theta_0)^2
 \tag{2.6}$$

Dihedral bond rotations are described by the angle χ between two plains, each containing coordinates of first and last three atoms for any

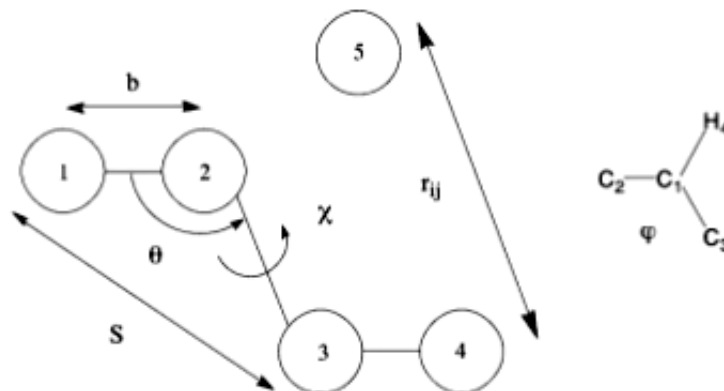


Figure 2.1: Notations of angles and planes accounted for an empirical force field.¹³

four covalently bound atoms, such as atoms 1, 2, 3 and 4 from Figure 2.1. If this angle can perform unrestricted rotation, it is called a proper dihedral motion. In the case of additional restraints, it is improper motion. Equation (2.7) gives the contribution due to the proper dihedral motions to the potential energy. In this equation, V is energy term specific for each type of the bond, n is the periodicity parameter, N_d is a constant which depends on the number of bonds and g is the phase parameter.

$$U_{dihedral} = \frac{V}{N_d}(1 + \cos(n\chi - g)) \quad (2.7)$$

Improper dihedral bond rotations are such motions of a dihedral angle which cannot fully rotate, depicted as the angle φ in 2.1. Equation (2.8) describes this type of motion, where φ and φ_0 are instantaneous and equilibrium improper angles, respectively.

$$U_{improper} = \frac{1}{2}k_\varphi(\varphi - \varphi_0)^2 \quad (2.8)$$

van der Waals interactions usually described as the Lennard-Jones (L-

J) 6-12 potential(Equation (2.9)). In this equation ε_{ij} is a L-J well-depth and σ_{ij} is a distance between atoms i and j where the potential is minimum.

$$U_{vdW} = 4\varepsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right) \quad (2.9)$$

Coulomb Interactions are electrostatic interactions between charged atoms and governed by the Coulombs law, as shown below:

$$U_{Coulomb} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}} \quad (2.10)$$

ϵ_0 is a relative permittivity of free space, while ϵ_r is a relative dielectric constant. q_i and q_j are charges of atoms i and j , respectively.

Numerical Integration

During MD simulations dynamics of all atoms in the system are defined by the Newtonian equations of motion. MD simulations are used to solve these equations of motion to obtain sets of trajectories in the phase space of the system (velocities and coordinates of all atoms). This is the basic information required for all further analysis of the system in question. To solve these equations, finite difference methods are commonly used. One of the most popular algorithms is the velocity Verlet method because it satisfies the conservation laws of energy and momentum and it is easy to implement numerically. Even though the algorithm has a global error proportional to Δt^2 , where Δt is a discrete time step, compared with Δt^5 for the Runge-Kutta algorithm, the velocity Verlet algorithm is computationally more efficient. The equations for advanced position and velocity for the velocity Verlet algorithm are given by Equations (2.11) and (2.12), respectively.

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \left(\frac{d\vec{r}_i}{dt}\right)_t \Delta t + \frac{1}{2}\left(\frac{d^2\vec{r}_i}{dt^2}\right)_t \Delta t^2 \quad (2.11)$$

$$\vec{v}_i(t + \Delta t) = \left(\frac{d\vec{r}_i}{dt}\right)_t + \frac{1}{2}\left(\left(\frac{d^2\vec{r}_i}{dt^2}\right)_t + \left(\frac{d^2\vec{r}_i}{dt^2}\right)_{t+\Delta t}\right) \Delta t \quad (2.12)$$

A general flow chart of MD simulations is illustrated in Figure 2.2.

According to this figure, initial coordinates and velocities of all atoms in the systems must be provided at the beginning of simulations. For biomolecules, coordinates can be retrieved from the Protein Data Bank (PDB).⁵⁹ Initial velocities can be sampled then from the Maxwell-Boltzmann distribution at particular temperature. These initial data are used for calculation of the potential energy given by Equation (2.2). Subsequently, the total potential energy is then used for calculation of the forces with Equation (2.3). Acceleration ($\vec{a} = \vec{F}/m$) can be calculated at this stage. Then positions and velocities at the next time step can be obtained from Equations (2.11) and (2.12) by solving equation of motion with the velocity Verlet algorithm. MD simulations will be stopped when the number of steps has reached a target number adequate to gain sufficient statistics of sampling the relevant region of the phase space of the system. It is important to note that the time interval Δt needs to be sufficiently small in order to be able to track atomic movements, for example Δt should be smaller than the time scale relevant to the fastest atomic movements being simulated. Typically, Δt is set to be close to 1 fs in MD simulations without adding the constraints associated with covalent bonds involving hydrogen atoms. However, if such bonds are constrained (usually with SHAKE⁶⁰ or RATTLE⁶¹ algorithms), Δt can be set to 2 fs to reduce computational time.

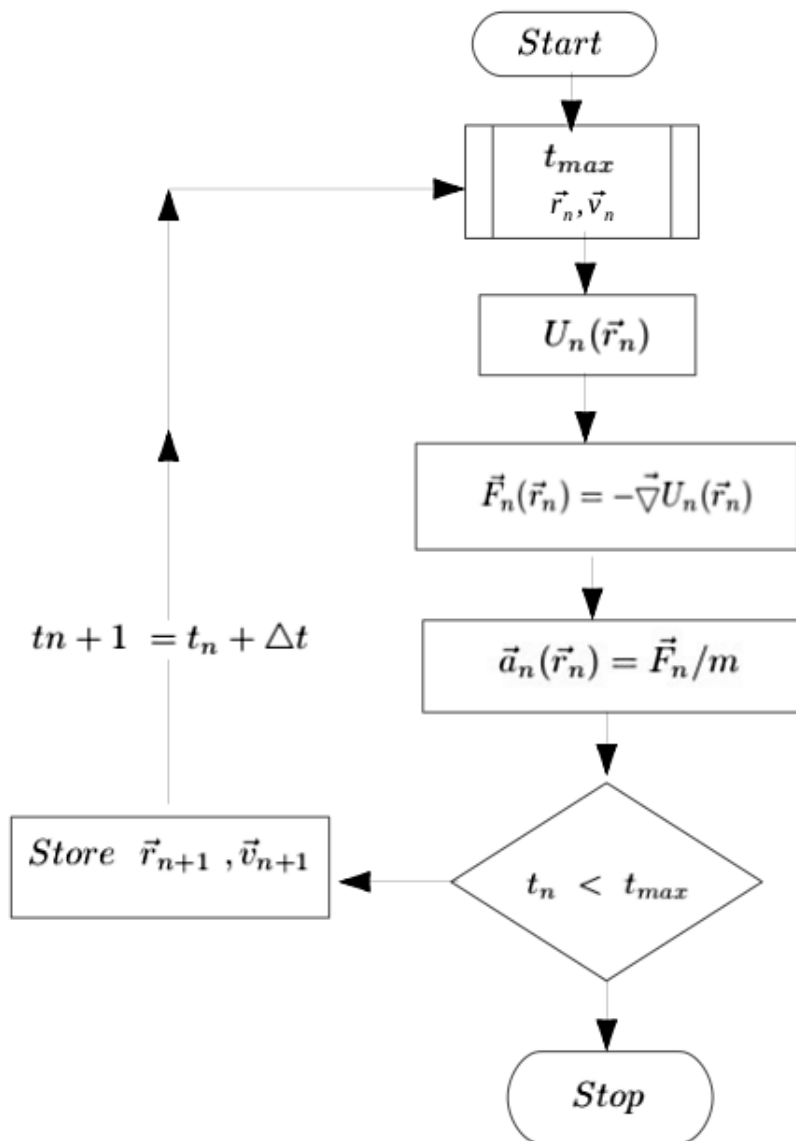


Figure 2.2: Molecular dynamics algorithm

Solvation Effects

Within the scope of this work, the aqueous environment that proteins reside in is extremely important and must be accounted for within the simulation of the system. Therefore, water molecules and ions must be accounted for. Typically, there are two types of solvation models in simulations, namely explicit and implicit solvation models. If water and ion molecules are represented in atomistic detail with a set of explicit molecules in the course of simulations, the approach is called an explicit solvation model. On the other hand, if water molecules and ions are simply treated as dielectric medium represented with parameters used to modify the interactions between protein atoms, this is an implicit solvation model. An example of the latter is the so-called Generalized Born continuum solvation model where the solvent is represented with a cavity in dielectric medium of certain shape recapitulating the dielectric screening resulting from solving the Poisson equation for the corresponding dielectric problem, and protein atoms are assigned with effective Born radii as a result.

The explicit solvation model can employ many types of water models, such as simple point charge (SPC)⁶², simple point charge extension (SPC/E)⁶², transferable intermolecular potential, three-position mode (TIP3P)⁶⁴ and transferable intermolecular potential, four-position mode (TIP4P))⁶⁵ models. Selection of a proper water model is important for accurate modeling of aqueous solutions. The explicit water models listed above are characterized by different charge distributions and geometry of water molecules. They also use different parameters for electrostatic (EL) and van der Waals (vdW) interactions such as partial charges and L-J parameters. For example, EL interactions for TIP3P, SPC, SPC/E water models are calculated based on the three charged points scheme, where two sites represent hydrogen and one site the oxygen atom. An additional (dummy) charged site is added along the H-O-H bisector between the hydrogens of a water molecule in TIP4P model; hence, the EL interactions are determined using four point charges. To select a water model for a given system, it is recommended to choose one that has already been optimized for a similar

system; usually one a force field developer has used. This is because, a force field for a specific system has been parameterized, using a specific water model. For example, CHARMM force field for biomolecules was parameterized with CHARMM version of the TIP3P water model.²⁵

In contrast to the explicit solvation model, implicit solvation models do not account for the atomistic structure of the solvent. Usually, implicit solvation models are less computationally expensive compared to all atom explicit solvent simulations. Moreover, solvation free energy arising from electrostatic, cavity formation and van der Waals interactions are calculated instantaneously, without the need for post-processing simulation trajectories. Two commonly used implicit solvation models are the Poisson - Boltzmann (PB) and generalized Born (GB) models with nonpolar effects taken into account with the solvent accessible surface area (SA) empirical parameterization, or the PB(GB)SA models. It is worth noting that in many cases implicit solvation models cannot describe the nonpolar terms of the solvation free energy with adequate accuracy,⁶⁷ which leads in turn to problems with structural solvation and hydrogen bonding between water molecules and a macromolecule.⁶⁷

Periodic Boundary Conditions

In MD simulations using explicit solvation models, the periodic boundary conditions (PBC) are usually used to minimize effects of finite simulation box size. In order to produce the periodic boundary conditions, a unit cell such as highlighted in Figure 2.3 will be replicated infinitely in two- or three- dimensions without walls inserted between the replicated cells. Figure 2.3 illustrates an example of two-dimensional periodic boundary conditions. The cells A, B, C, D, E, F, G and H are images of the highlighted cell. Another advantage of the periodic boundary conditions is that they keep the number of atoms in real space (highlighted box) constant. As seen in Figure 2.3, while atom 1 is leaving the box in real space, its image 1_G is entering at the opposite side of the box. Nevertheless, the periodic boundary conditions also cause some problems with the calculations of vdW and EL interactions. Since cells in the periodic boundary condi-

tions span infinitely, as a consequence a course of simulations may take a long time to calculate long range interactions like vdW and Coulomb interactions. Special treatments for these interactions include the use of cutoff distances, where vdW interactions are truncated. However, this technique is not useful for calculating EL interactions. This is because the EL interactions are long range; their magnitudes decay with distance as r^{-1} , compared to r^{-6} in the case of vdW interactions. If the cutoff is used, calculations of the EL interactions may result in a huge error. Thus, a new method called particle-mesh-Ewald (PME) summation is employed for EL interactions calculated from the sum of EL in all unit cells with special treatments (please see ref^{13,14}).

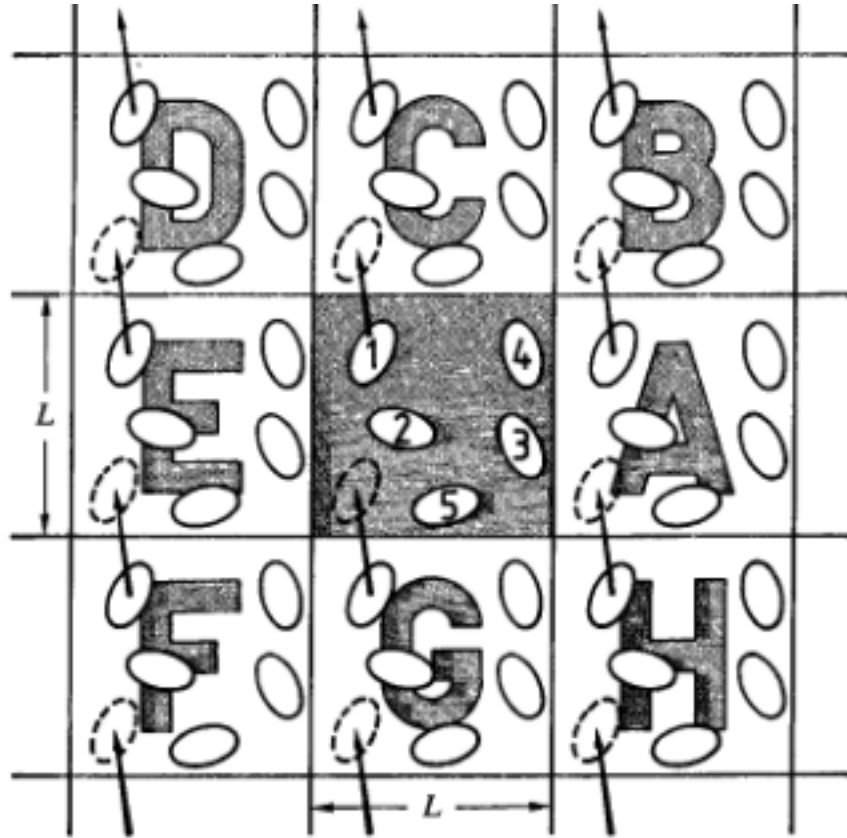


Figure 2.3: Periodic boundary condition(PBC)¹⁴

Canonical (NTV) and Isothermal-isobaric (NTP) ensembles

In order to carry out simulations at a fixed temperature, a system is generally coupled to a reservoir so as to control the temperature. Therefore, additional terms should be added to Newtons equation of motion to account for coupling between a system and the thermostat. In calculations presented in this thesis, the Langevin dynamics will be employed to control temperature of the system. Hence, the modified Newtons equation (the Langevin equation, see Equation (2.13)) is used to generate Boltzmann distribution for NTV ensemble .

$$M_i \frac{d\vec{v}(\vec{r}_i)}{dt} = \vec{F}_i(\vec{r}_i) - \gamma \vec{v}_i(\vec{r}_i) + \sqrt{\frac{2\gamma k_B T}{M_i}} R(t) \quad (2.13)$$

where (2.13), M_i is mass of atom i , $\vec{F}_i(\vec{r}_i)$ is the force acting on atom i , \vec{r}_i is its position, k_B is the Boltzmann constant, T is the temperature, and $R(t)$ is a univariate Gaussian random process. The last 2 terms in Equation (2.13) are from coupling between the system and the reservoir. The equation is then integrated with the Verlet method as discussed earlier.

In case of NTP ensemble, the idea of controlling pressure of the system in MD is to consider a cell in periodic boundary conditions as a container. The volume of the cell is allowed to change without changing shape of the cell to adjust the pressure to get the specified value. The internal pressure is then related with change in volume of the cell as seen in Equation (2.14) below.¹⁸

$$P_{int}dV = \sum_i (m_i \vec{v}_i^2(\vec{r}_i) + \vec{F}_i(\vec{r}_i)) - dV \frac{\partial U(V)}{\partial V} \quad (2.14)$$

where , P_{int} is the internal pressure of the system, V is its volume, $U(V)$ is the potential energy. Here m_i , \vec{r}_i , $\vec{v}_i(\vec{r}_i)$ and \vec{F}_i are mass, position vector, velocity and forces acting on the atom i of the system, respectively. Equation (2.14) is then solved using the velocity Verlet algorithm following as a part of MD algorithm represented in Figure 2.2.

2.2.3 Molecular Mechanics Concepts

Generally, MM uses classical mechanics to obtain an equilibrium structure of the system by minimizing the systems molecular potential energy surface. MM algorithms may be vaguely divided into two groups, algorithms with and without use of derivatives with respect to the coordinates of a structure. In this section, only the first-order derivative methods, in particular, the steepest descent and the conjugate gradient methods, will be briefly discussed (details can be found elsewhere).²

The steepest decents method searches for minimum energy by moving atoms so as to reduce potential energy (Equation (2.4)) closer to the minimum point in the direction where the potential energy decreases most quickly. This is the direction of $-\vec{\nabla}U(\vec{x})$, where $U(\vec{x})$ is potential energy. If the search starts at an arbitrary point \vec{x}_k and move to the next point \vec{x}_{k+1} through Equation (2.15), where \vec{g}_k and λ_k are gradient and step length at point \vec{x}_k , respectively.

$$\vec{x}_{k+1} = \vec{x}_k - \lambda_k \vec{\nabla}U(x) = \vec{x}_k - \lambda_k \vec{g}_k \quad (2.15)$$

To move to point \vec{x}_{k+1} where $U(\vec{x}_{k+1})$ is minimum ($\frac{dU(x_{k+1})}{d\lambda_k} = 0$, more details elsewhere⁶⁸), λ_k is chosen where $\vec{\nabla}U(\vec{x}_{k+1})$ and \vec{g}_k are orthogonal. The next step is then taken in the direction (\vec{d}_{k+1}) of $-\vec{\nabla}U(\vec{x}_{k+1})$. The algorithm of the steepest decents method is shown in Table 2.1 below. Due to gradient at a current point and direction for the next point are necessary to be orthogonal; therefore, this method produces a zig-zag pattern especially when moving along a narrow valley as approaching to minimum potential energy as seen in Figure 2.4.

Table 2.1: Method of steepest decent⁶⁸

Initializing:	$\vec{g}_k = \vec{\nabla}U(\vec{x}_k), \vec{d}_k = -\vec{g}_k$
Determine the step length λ_k :	$\min U(\vec{x}_k + \lambda_k \vec{d}_k), \text{ where } \lambda_k > 0$
Calculate the new point:	$\vec{x}_{k+1} = \vec{x}_k + \lambda_k \vec{d}_k.$
Calculate gradient:	$\vec{g}_{k+1} = \vec{\nabla}U(\vec{x}_{k+1}).$
Set direction of search:	$\vec{d}_{k+1} = -\vec{g}_{k+1}$

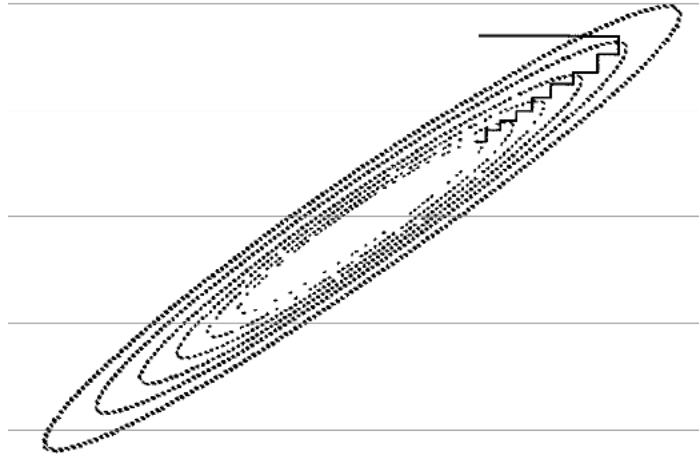


Figure 2.4: Oscillation or zig-zag patterns of move in a long narrow valley shown as contour lines.²

Conjugate gradients minimisation is also applied to a system by moving atoms in the system to reduce potential energy which oscillation during moving along a narrow valley is not shown. This is because direction at the next point and gradient at the previous point in this method are not necessarily orthogonal while choosing step length. The directions are Q orthogonal and linearly independent, where Q is a symmetric positive definite matrix. If the search starts at point \vec{x}_k , to move to the point \vec{x}_{k+1} via Equation (2.15), the step length λ_k is determined where $U(x_{k+1})$ is minimum by Equation (2.16). In Equation (2.16), \vec{d}_k , \vec{d}_k^T , \vec{g}_k , \vec{g}_k^T and \vec{Q} are direction, transpose of direction, gradient, transpose of gradient at

point \vec{x}_k and symmetric positive definite matrix, respectively.

$$\lambda_k = \frac{\vec{g}_k^T \cdot \vec{g}_k}{\vec{d}_k^T \cdot (\vec{Q} \cdot \vec{d}_k)} \quad (2.16)$$

The direction at next point is defined as Equation (2.17)

$$\vec{d}_{k+1} = -\vec{g}_{k+1} + \gamma_k \vec{d}_k \quad (2.17)$$

In Equation (2.17), γ_k is defined below.

$$\gamma_k = \frac{\vec{g}_{k+1}^T \cdot \vec{g}_{k+1}}{\vec{g}_k^T \cdot \vec{g}_k} \quad (2.18)$$

The algorithm of the conjugate gradient method is shown in Table 2.2

Table 2.2: Method of conjugate gradient⁶⁹

Initializing:	$\vec{g}_k = -\vec{\nabla}U(\vec{x}_k), \vec{d}_k = -\vec{g}_k$
Determine the step length λ_k :	$\min U(\vec{x}_k + \lambda_k \vec{d}_k), \text{ where } \lambda_k > 0$
Calculate the new point:	$\vec{x}_{k+1} = \vec{x}_k + \lambda_k \vec{d}_k.$
Calculate gradient:	$\vec{g}_{k+1} = -\vec{\nabla}U(\vec{x}_{k+1}).$
Set direction of search:	$\vec{d}_{k+1} = -\vec{g}_{k+1} + \gamma_k \vec{d}_k, \text{ where } \gamma_k = \frac{\vec{g}_{k+1}^T \cdot \vec{g}_{k+1}}{\vec{g}_k^T \cdot \vec{g}_k}$

2.3 Summary

This chapter addresses the thermodynamics of protein adsorption, an overview of MD simulations, MD methods emphasized on standard force field equations, which are a combination of bonded and non-bonded interactions, numerical integration often used for solving Newtonian equation, particularly the Verlet algorithm, solvation models and use of periodic boundary conditions in MD simulation. The chapter also briefly discusses MM simulations which are usually conducted to find optimum structure of the system before performing long time MD simulations.

Chapter 3

Molecular Dynamic Simulations of Human Lysozyme and Apo- α -Lactalbumin

3.1 Introduction

Lysozyme and α -lactalbumin were employed as model proteins for understanding how charge and stiffness of the protein affects their adsorption to PEGylated surfaces. These were chosen primarily because both proteins have approximately the same shape, size, folding pattern and number of disulfide bonds. Despite these similarities, their functions are very different: lysozyme hydrolyses glycosidic bonds of sugar molecules, while α -lactalbumin takes part in lactose synthase.³⁵ Also, α -lactalbumin is capable of binding Ca^{2+} ions at two sites, whereas lysozyme cannot. Only Ca^{2+} ion bounded at the primary site of α -lactalbumin plays an important role in its structural integrity.³⁵ In the absence of Ca^{2+} ions, α -lactalbumin is referred to as apo- α -lactalbumin, while in the presence of Ca^{2+} ions the protein is called holo - α -lactalbumin. In this chapter, properties of these proteins in solution will be investigated by performing Molecular Dynamics

(MD) simulations.

3.2 Preliminary studies of X-ray crystal structure of Human lysozyme and α -Lactalbumin

Before proceeding with modeling of lysozyme and α -lactalbumin in physiologically relevant solutions, folding patterns, stabilities and solvent accessible surfaces of the proteins were analyzed based on their X-ray structures, by using the Visual Molecular Dynamics (VMD) software. Then, in the next subsection, MD simulations were carried out to reveal the structural stability of the proteins under physiological conditions. The crystal structures of human lysozyme (PDB accession code 1REX)³³ and α -lactalbumin (PDB accession code 1A4V)³² were taken from the Protein Data Bank (PDB).

Chain length, Shape and Secondary structure

Lysozyme is composed of 130 residues, with an ellipsoid shape of $4.5 \times 3.0 \times 3.0$ nm (Figure 3.1a) as determined using X-ray crystallography techniques with a 1.50 Å resolution.³³ Similarly, based on 1.8 Å resolution X-ray crystal structure, α -lactalbumin³² is also an ellipsoid protein (Figure 3.2b) composed of 123 residues with a dimension of $2.5 \times 3.7 \times 3.2$ nm. Due to having nearly the same number of residues, the molecular weights of both proteins are similar at approximately 14.4 and 14.2 kDa for lysozyme and α -lactalbumin, respectively. Additionally, both proteins also share similar folding patterns, which are composed mainly of α -helices and β sheets as seen in Figure 3.1b and Figure 3.2b, for lysozyme and α -lactalbumin, respectively. In both cases, the molecular structures are characterized by grouping of α -helices and β sheets in domains, with the α domain includes residues 5-36 and 90-115 comprising α 1 (5-14), α 2 (25-36), α 3 (90-100) and α 4 (110-115). The β domain is located at residues 41-60 comprising β 1 (43-46), β 2 (51-54) and β 3 (59-60). In the cartoon representation of Figure 3.1b, α helices and helix 3/10 are shown in purple and blue colours.

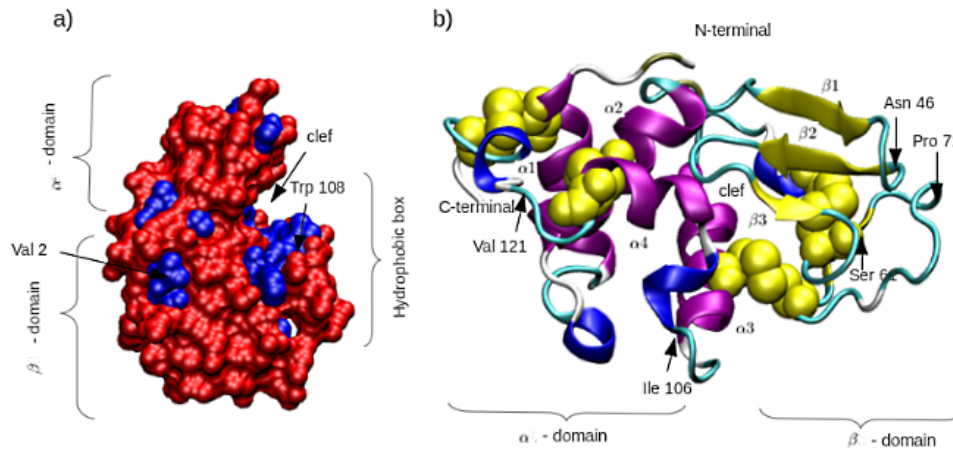


Figure 3.1: a) Shape of lysozyme. Solvent accessible surfaces corresponding to hydrophilic and hydrophobic residues are shown in red and blue, respectively. b) Secondary structures of lysozyme; 4 α helices ($\alpha 1, \alpha 2, \alpha 3, \alpha 4$), 4 helices 3/10, β sheets ($\beta 1, \beta 2$ and $\beta 3$), turns and coils depicted as purple and blue and yellow, cyan and white colours, respectively, cystein residues illustrated as yellow van der Waals (vdW) spheres. Residues marked by arrows are highly flexible residues. (indicated as white arrows in Figure 3.3 a).

While turns, coils and main- β sheets are depicted as cyan, white colours and yellow arrows, respectively. In case of α - lactalbumin, its α domain is found in residues 1-36 and 86-111, comprising of $\alpha 1$ (5-11), $\alpha 2$ (23-34), $\alpha 3$ (86-98) and $\alpha 4$ (106-110), whereas the β domain of this protein is found in residues 41-56, composed of $\beta 1$ (41-43), $\beta 2$ (48-50) and $\beta 3$ (55-56) as seen in 3.2b. α helices and helices 3/10 of the protein are still depicted as purple and blue helical cartoons, while turns, coils and main- β sheets are in cyan, white rods and yellow arrows, respectively. Main features of the secondary structures of both proteins and their locations are summarized in Table 3.1.

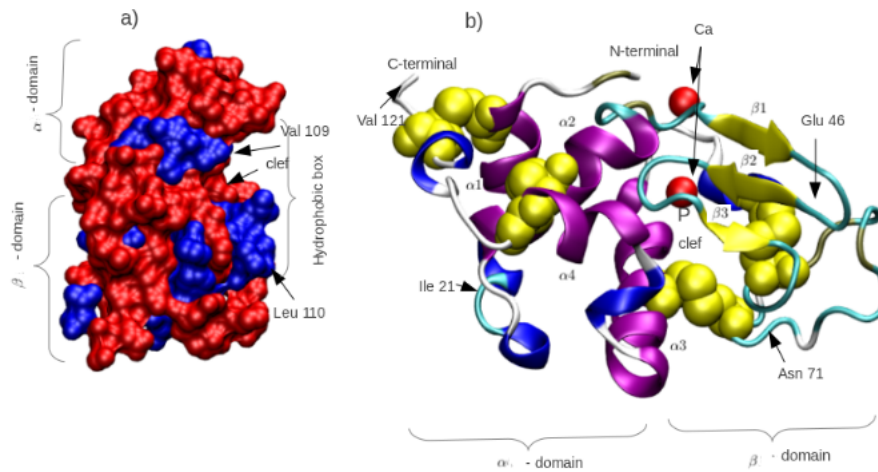


Figure 3.2: a) Shape of α -lactalbumin. Solvent accessible surfaces corresponding to hydrophilic and hydrophobic residues are shown in red and blue colours, respectively. b) Secondary structures of holo- α -lactalbumin; 4 α helices (α_1 , α_2 , α_3 , α_4), 5 helices 3/10, β sheets (β_1 , β_2 and β_3), turns and coils depicted as purple, blue, yellow, cyan and white colours respectively, cystein residues illustrated as yellow van der Waals (vdW) spheres. Residues marked by arrows are highly flexible residues. (indicated as white arrows in Figure 3.3 b).

Table 3.1: Secondary structure of human lysozyme and human holo- α -lactalbumin assigned using the VMD software.

Regions	Proteins (Residue number)	
	Lysozyme	α -lactalbumin
$\alpha 1$	5-14	5-11
$\alpha 2$	25-36	23-34
$\beta 1$	43-46	41-43
$\beta 2$	51-54	48-50
$\beta 3$	59-60	55-56
Helix 3/10	20-21, 81-85, 105-108, 121-124	13-16, 18-21, 76-81, 102-105, 117-120
$\alpha 3$	90-100	86-98
$\alpha 4$	110-115	106-110

Disulfide Bridges

The number of disulfide bonds formed in both lysozyme and α -lactalbumin molecules are comparable. In lysozyme molecule, there are four disulfide bonds formed between residues Cys 6, Cys 128, Cys 30- Cys 116, Cys 65- Cys 81 and Cys 77- Cys 95 as illustrated with black arrows linking between the above Cys pairs in Figure 3.3 a. Similarly, in α -lactalbumin there are 4 disulfide bridges formed between residues Cys 6- Cys 120, Cys 28- Cys 111, Cys 61- Cys 77, and Cys 73- Cys 91 as drawn in black arrows linking between the corresponding pairs of the Cys residues. It is interesting that the positions of disulfide bridges are conserved in both proteins.

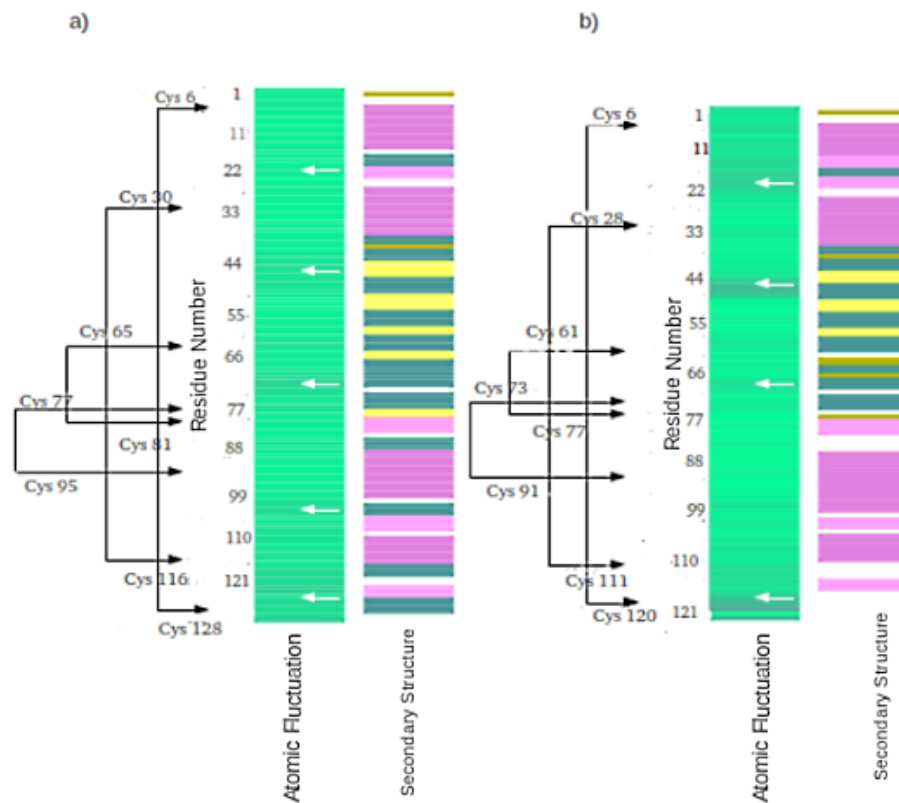


Figure 3.3: Atomic fluctuation, secondary structure of each residue and disulfide bridges formed in lysozyme a) and holo- α -lactalbumin b) labeled as arrows linking cysteine residues. Different values of temperature β -factor are illustrated by different intensity of colour, with darker colour indicating higher values, and consequently higher flexibility. Domains with noticeably high fluctuations marked with white arrows. In secondary structure maps, α helices, helix 3_{10} , main β sheets, partial- β sheets, turns and coils are depicted with purple, pink, yellow, gold, green and white, respectively.

Solvent Accessible Surfaces

Generally, hydrophobic residues (Leu, Val, Ile, Ala, Gly, Phe, Tyr, Trp, Met, and Pro) are buried inside of a protein, while hydrophilic residues (Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Lys, Arg, His, Glu and Asp) are exposed to the solvent. However, according to Figure 3.1a, some hydrophobic residues of lysozyme (such as Val 2 and Trp 108 illustrated as blue surface) are solvent accessible. Similarly, in case of holo- α -lactalbumin, some hydrophobic residues are also exposed to solvent even more than that of lysozyme, as depicted in blue surfaces in Figure 3.1a. According to Figure 3.1 and Figure 3.2, interestingly most hydrophobic residues of both proteins are located between α and β domains. Therefore, this area is called hydrophobic box as indicated in Figure 3.1a and Figure 3.2a for lysozyme and α -lactalbumin, respectively.³⁵ On the other hand, solvent accessible surface corresponding to hydrophilic residues of both lysozyme and α -lactalbumin are drawn as red in Figure 3.1.

Figure 3.4a and b shows charts indicating values of solvent accessible surfaces of lysozyme and holo- α -lactalbumin, respectively; darker colours indicate more buried surfaces, white colour presents 100 % solvent accessible surfaces. Regarding to Figure 3.4a and b, the residues corresponding to helix 3/10 lying close to C- terminal of both proteins (Residues 121-124 and residues 117-120 in lysozyme and holo- α -lactalbumin, respectively.) are more exposed to solvent than residues forming α helices and β sheets. Moreover, some residues corresponding to turn and coil conformations such as residues 106 and 61 in lysozyme molecule and residues 21 and 71 in holo- α -lactalbumin molecule are also highly exposed to the solvent. Interestingly, the residues which are highly exposed to solvent (i.e. helices 3/10 lying close to C-terminal, some portions of turns and coils) are also highly flexible as indicated by the darker colours in the atomic fluctuation (β factors) portion of Figure 3.3 a and b for lysozyme and holo- α - lactalbumin, respectively especially, in the areas marked with white arrows which are mostly turns, coils and helices 3/10. Information of β factors corresponding to secondary structures for others residues is shown in Figure 3.3 a and

b for lysozyme and holo- α -lactalbumin, respectively.

Residues in Catalysis

Due to difference in amino acid sequences between lysozyme and α -lactalbumin, their functions are different.³⁵ Lysozyme binds sugar and hydrolyzes its glycosidic bonds at the sites located in residues 101, 103, 107, 108, 109, 34, 35, 37, 53, 57, while α -lactalbumin involves in lactose synthase which takes place in residues 31, 32, 110 and 118.³⁵ It is noted that these active sites are in the cleft region of both proteins.³⁵

Salt Bridges

Regarding to the X-Ray crystal structure of lysozyme³³ and α -lactalbumin³², only lysozyme has three salt bridges formed between residues Asp 102 - Arg 98, Glu 7 - Lys 1 and Asp 120 - Arg 122,³³ while no salt bridges are observed in α -lactalbumin molecule.³² A protein with higher number of salt bridges formed may show higher stability compared to a similar protein (e.g., having similar number of residues and secondary and tertiary structure) with fewer or no salt bridges.

Ligands for Ca²⁺ Binding

Human lysozyme is not capable of binding Ca²⁺ ion³⁵, while only human α -lactalbumin which is considered as a Ca²⁺ bounded protein at 2 different sites namely primary and secondary sites but lysozyme.³² The primary Ca²⁺ binding site is located between residues 77-80 (helix 3/10) and residues 86-98 (α 3) as shown in Figure 3.2 b with red vdW spheres lying close to the character P. Besides, it was reported that at the Ca²⁺ ion at this site plays an important role in its structural stability, and in the absence of the Ca²⁺ bounded at this site its stability decreased considerably^{35,39}. However, Ca²⁺ ion bounded at the secondary site of the protein does not play any roles in structural stability of the proteins.^{35,39}

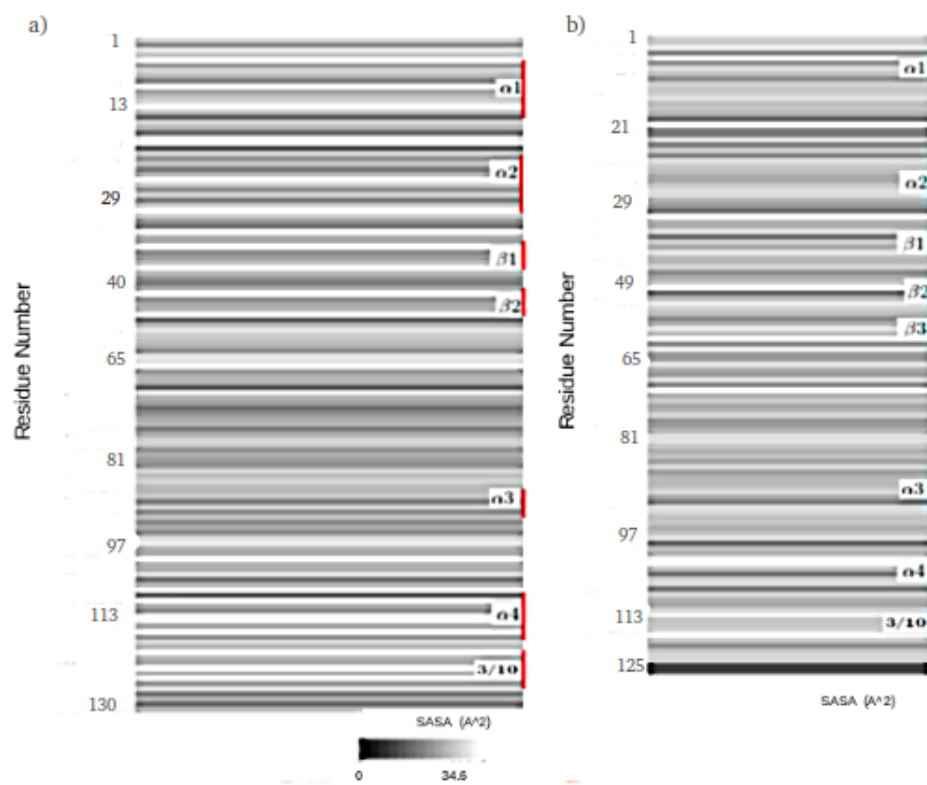


Figure 3.4: Charts representing values of solvent accessible surfaces of lysozyme a)³³ and holo- α -lactalbumin b)³², different scales of colour indicating different percentage of solvent accessible surfaces: white colour indicating 100 % solvent accessible surfaces, while black colour indicating totally buried surfaces.

Basic and Acidic Groups

Lysozyme is considered as a basic protein, while α -lactalbumin is classified as an acidic protein. This is because the total number of acidic and basic residues of both proteins are different. α -lactalbumin composes of 20 acidic residues (12 Asp, 8 Glu) and 13 basic residues (2 His, 12 Lys, 1 Arg). In contrast, lysozyme consists of 11 acidic residues (8 Asp, 3 Glu), and 20 basic residues (1 His, 5 Lys and 14 Arg). As a result of higher total number of acidic residues than basic ones, α -lactalbumin is called the acidic protein with PI of around 4.5 and with charge of -7e at physiological pH.³⁵ While the total number of basic residues in lysozyme exceeds that of acidic residues; thus, it is named as the basic protein with high PI of approximately 11 and with charge of +8e under physiological pH.⁴⁰

3.3 Structural Analysis of Human Lysozyme and Human Apo- α -Lactalbumin Using Molecular Dynamics Simulations

3.3.1 Methodology

Model and Simulations Description

The initial PDB crystal structure of the wild-type human lysozyme (1REX.pdb)³³ and α -lactalbumin (1A4V)³² were utilized for all modeling aspects. The pKa value for each residue of both proteins were determined with PROPKA 3.1 software. This information then was used to assign charged states of protonatable residues based for pH values of 7.0. Resulting structure was used as input for molecular dynamics simulations with Amber 11 molecular dynamics package. Predicted pKa values of all Asp and almost all Glu residues of lysozyme fall between 3.24 - 4.75 (see Appendix A); therefore, they were treated as deprotonated residues. Even though pKa value of Glu 35 lying in non-polar environment³⁵ (7.31) is slightly higher than pH at simulation conditions, as a result it was still treated as deprotonated

residue.³⁸ All basic residues, namely Lys, Arg and His were predicted to be protonated. In the case of α -lactalbumin, all acidic residues were treated as being deprotonated, while all basic residues were treated as protonated. pKa value of Asp 88 is 7.25 (see appendix A) is just slightly higher than pH at simulation condition. Therefore, in our simulations Asp 88 of α -lactalbumin was treated as deprotonated. All water molecules and ions (including Ca^{+2} ions in the α -lactalbumin molecule) were removed. Subsequently, the modified X-Ray crystal structure of lysozyme was solvated in $18.77 \text{ \AA} \times 18.77 \text{ \AA} \times 18.77 \text{ \AA}$ periodic water box with 12807 TIP3P water molecules. Likewise, the modified X-ray crystal structure of α -lactalbumin was also solvated in $18.77 \text{ \AA} \times 18.77 \text{ \AA} \times 18.77 \text{ \AA}$ periodic box with 11502 TIP3P water molecules. Then, 8 Cl^- and 7 Na^+ counter ions were added with the AMBER package to neutralize lysozyme and apo- α -lactalbumin systems, respectively.

Two minimization runs then were carried out for both systems. The first run was to relax solvent and ions molecules with 5000 steps of the steepest decents method followed by 5000 steps of the conjugate gradient method. The second run was performed with the same methods and number of steps as the first run (5000 steps of steepest decents method followed by 5000 steps of conjugate gradient method) without restrain. Next, the entire solvated system was thermalized by gradually increasing temperature from 0 to 310.15 K over 100000 MD steps. The temperature regulation was controlled by Langevin dynamics, with collision frequency of 5 ps^{-1} for all MD simulations. The systems were then subject to another 150000 MD steps at constant temperature of 310.15 K and constant pressure of 1 atm to relax solvent molecules and to adjust density of the system to 1 g/cm^3 . The relaxation time for pressure regulation was 1 ps^{-1} . The time step used during thermalization and equilibration was set to 1 fs. Before proceeding to production runs, another 50000 MD steps were performed with the SHAKE⁶⁰ algorithm applied to all bonds involving hydrogen atoms to remove high frequency modes (i.e. C-H stretches) from a system to allow the use of a larger time step. Hence, the time step was set to 2 fs. The relaxation time for the pressure regulation was set to 2 ps^{-1} . Before proceeding

to long production runs, the convergence study was conducted by performing 100000 MD steps with a solvated lysozyme system with different number of CPUs at Westgrid⁷⁰(<http://www.westgrid.ca>). An optimum number of CPUs was then used for subsequent long production runs. In the production runs, all parameters were set to the same values as used during pre-production runs (with the number of steps extended to 600000). During the production runs coordinates of the systems were saved every 0.2 ps. In all simulations, the Amber ff03.r1 force field was used. The particle-mesh Ewald (P3MP) algorithm was applied for treatments of long ranged electrostatic (EL) interaction. The cut off value of 16 Å was set for truncation of vdW interactions during simulations.

Analysis methods

Root mean square deviation (RMSD) is calculated via Equation (3.1), where N is total number of atoms \vec{r}_{1i} and \vec{r}_{2i} are initial and final position vectors of atom i .

$$RMSD = \sqrt{\frac{\sum_{i=1}^N (\vec{r}_{1i} - \vec{r}_{2i})^2}{N}} \quad (3.1)$$

Root mean square fluctuation (RMSF) can be determined via Equation (3.2), where T is total number of trajectory frames, \vec{r}_t and $\langle \vec{r} \rangle$ are instantaneous (i.e. trajectory frame t) and average position vectors, respectively.

$$RMSF = \sqrt{\frac{\sum_{t=1}^T (\vec{r}_t - \langle \vec{r} \rangle)^2}{T}} \quad (3.2)$$

Radius of gyration (R_g) is defined as Equation (3.3) below, where \vec{R}_c and \vec{r}_i are position vectors of a protein center of mass and that of the atom i , respectively. M and m_i are total mass of a protein and mass of atom i , respectively.

$$R_g = \sqrt{\frac{\sum m_i (\vec{r}_i - \vec{R}_c)^2}{M}} \quad (3.3)$$

3.3.2 Results and Discussion

Scaling of MD calculations with number of CPUs

Figure 3.5 shows time require to perform all-atom explicit solvent MD simulations for lysozyme (10000 MD steps) as a function of number CPUs. According to Figure 3.5, the simulation time is reduced remarkably as number of CPUs increases. However, the simulation time remains almost constant when more 32 CPUs were used. Therefore, further simulations were performed by using 32 CPUs to optimize usage of computational resources.

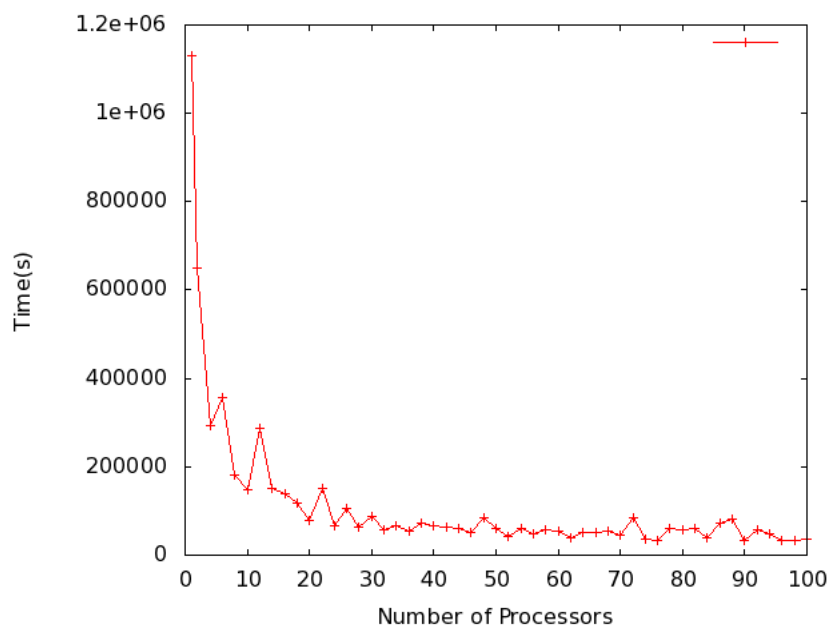


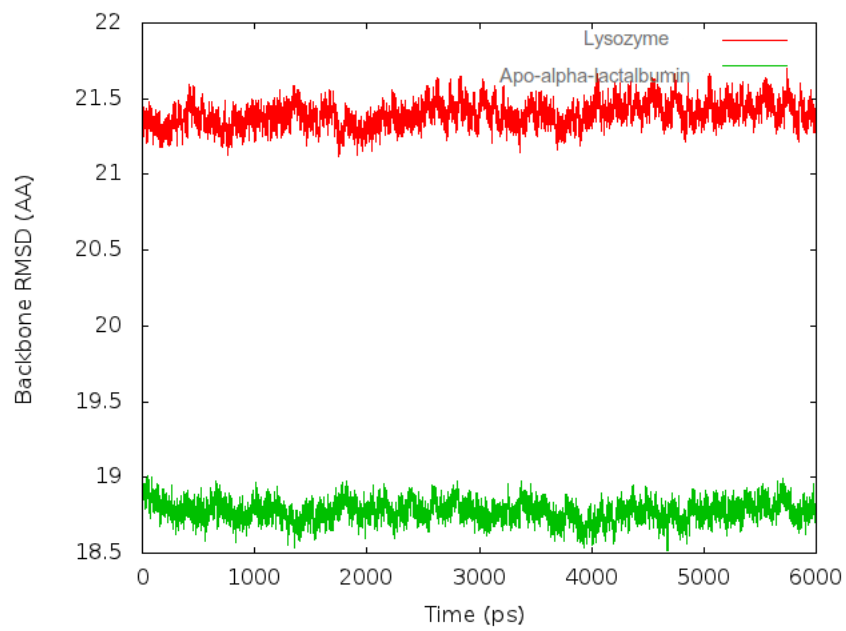
Figure 3.5: Time required to carry out 1 ns MD simulations for human lysozyme as a function of number of CPUs used. 32 is optimum number of CPUs for this system.

Analysis of root mean square deviation/fluctuation (RMSD/RMSF)

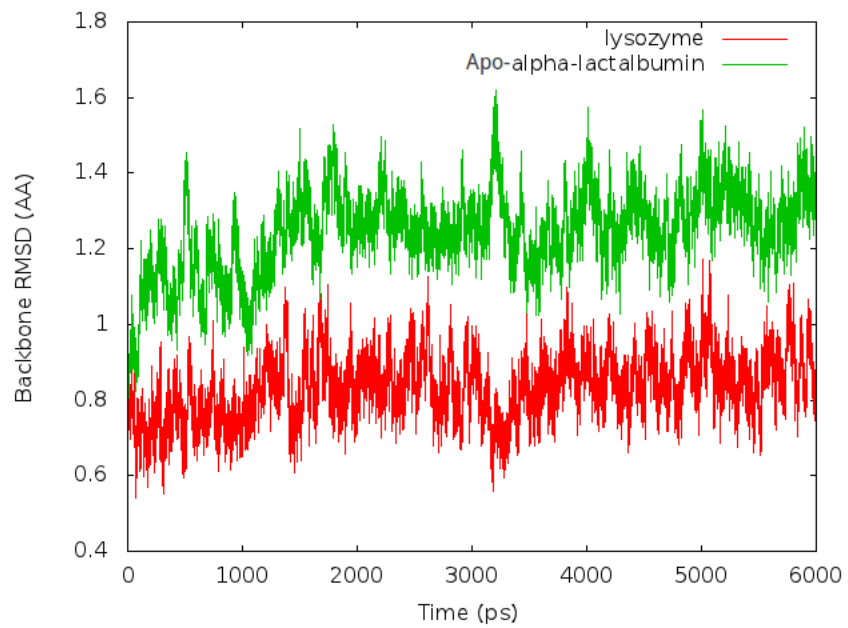
Simulations of lysozyme and α -lactalbumin in an absence of Ca^{2+} ions (apo- α -lactalbumin) have been performed under physiological condition (pH 7.0). Figure 3.6(a) shows a comparison of backbone RMSD of human lysozyme and apo- α -lactalbumin relative to their crystal structures. Overall, the backbone RMSDs of both proteins saturate at large constant values, around 21.4 and 18.75 Å for lysozyme and apo- α -lactalbumin, respectively. Large values of RMSD indicate changes of the tertiary and secondary structures of the proteins introduced into simulation environments. A RMSD value larger than 10 Å with respect to the crystal structure, was also found in a previous study with the wild type human lysozyme submerged into 8M urea water solution at pH of 7.0 and at temperature of 300 K.³⁸ In spite of large RMSD observed in this study, the secondary structures of the protein were stable over 15 ns. The provided explanation was not clear. However, they mentioned that RMSD might not be a useful tool for the local structural analysis when its value is large. Due to the fact that lysozyme undergoes an extreme pH change from 4.2 to 7.0, while α -lactalbumin moves only from pH 6.5 to 7.0, this may be the reason for a higher deviation of backbone positions in case of lysozyme. This is because both proteins are trying to optimize their structure under new environments. Interestingly, we found relatively low values of the backbone RMSD calculated with respect to the minimized structures for each protein, see Figure 3.6(b). The backbone RMSD of apo- α -lactalbumin is slightly higher than that of lysozyme, at approximately 1.4 and 1 Å for apo- α -lactalbumin and lysozyme, respectively. This indicates that secondary and tertiary structures of both proteins should not be denatured because of only slightly change in backbone RMSD relative to minimized structure noticed over the course of simulations.

However, as a result of the unreliable-backbone RMSD (i.e. Larger and smaller backbone RMSD of lysozyme than that of apo- α -lactalbumin relative to X-Ray crystal structure and minimized structure of each protein, respectively.), the backbone RMSF has been calculated and illustrated in

Figure 3.7. In general, the backbone RMSF profiles of both proteins are quite similar with higher fluctuations found at solvent exposed residues, especially turns and coils, with higher RMSFs for α -lactalbumin. The comparable backbone RMSF profiles of both proteins could be a result of similarity in their folding patterns which might be due to 30-45 % sequence similarity of both proteins.³⁵



(a)



(b)

Figure 3.6: a) Backbone RMSD with respect to crystal structure of each protein, b) Backbone RMSD with respect to structure of each protein after minimization runs.

Regarding to Figure 3.7, peak positions of the backbone RMSF of both proteins are mostly located in turn and helix 3/10 structures as drawn in green colour in the pictures showing secondary structures of lysozyme (right) and apo- α -lactalbumin (left) in the top of Figure 3.7, which are exposed to solvent under X-Ray crystal environments (Figure 3.4). It is noticeable that highly flexible residues of both proteins under crystal environments (Residues marked as white arrows in Figure 3.3a and b, for lysosyme and α -lactalbumin, respectively.) also show high flexibility in simulations. According to Figure 3.7, $\alpha 2$ is the most stable element of the secondary structure for both proteins. This is because this region is least exposed to solvent, as a consequence it is least perturbed by frictions from solvent which may induce fluctuation of an exposed residue. Additionally, this region is surrounded by other conformations; thus, it is difficult to fluctuate. Also, this region is clamped by a disulfide bond between residues 30-116 in lysozyme molecule and residues 28-111 in α -lactalbumin molecule. Thus, it seems reasonable that this region is more stable than other regions. It is also noticeable in Figure 3.7 that backbone RMSF at N- and C- terminal of both proteins are remarkably high. It has been reported that motions of main conformations also induce (i.e. α - helices and β - sheets) increase in flexibility of N- and C-terminal of both proteins.³⁷ Moreover, extremely high fluctuation in C-terminal of apo- α -lactalbumin was reported to be a result of the absence of Ca^{2+} ion bounded at its primary site.^{37,39} According to these studies^{37,39}, in the absence of Ca^{2+} ion mobility of main chain and side chain at Ca^{2+} binding site (between residue 79-88 and 86-98) increased more than when the protein bonded to Ca^{2+} ion; hence, this induced C-terminal to fluctuate more.

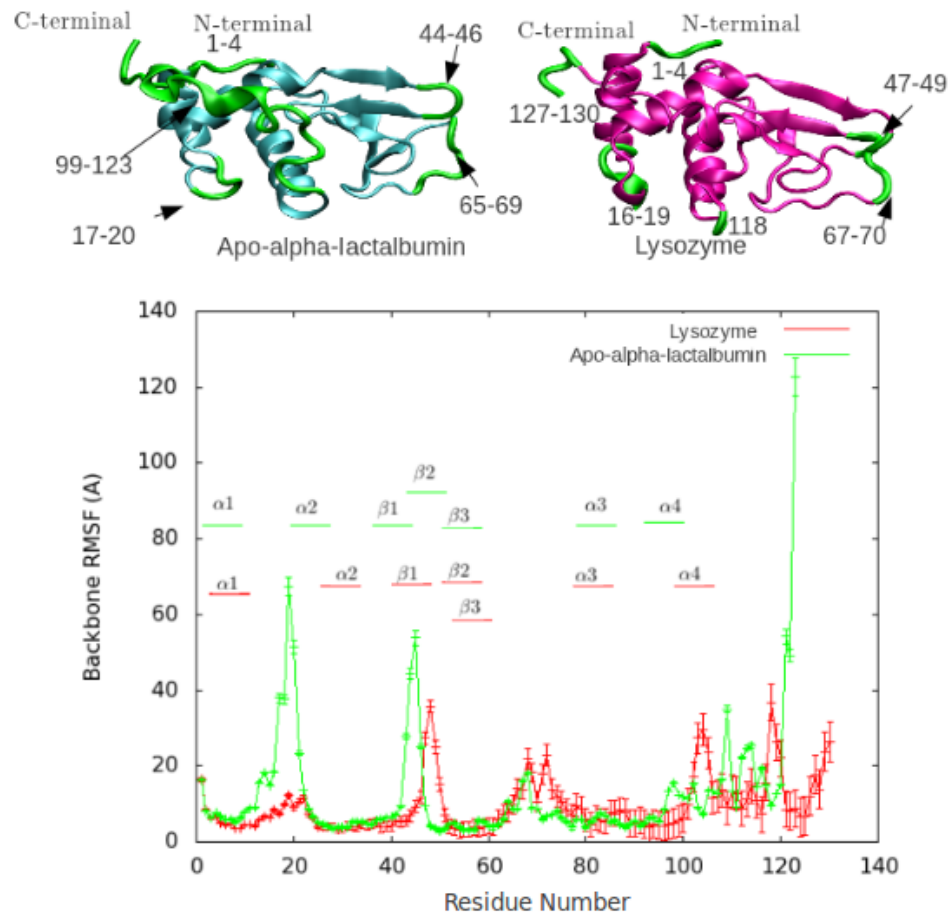


Figure 3.7: Backbone RMSF with error bars of lysozyme and apo- α -lactalbumin. Regions with highly fluctuating residues are depicted as green colour in cartoon representations of lysozyme (top right) and apo- α -lactalbumin (top left).

Analysis of Secondary Structure

Despite the high backbone RMSDs relative to the X-Ray crystal structures, the secondary structure is mostly preserved in the course of MD simulations for both proteins, as seen in 3.8 and Figure 3.9 for lysozyme and apo- α -lactalbumin, respectively. However, it is noticeable that lysozyme is more stable than apo- α -lactalbumin throughout the course of simulations. In case of lysozyme only a slight secondary transition from turns to coils in residues 117-120 lying at a loop connecting α 4 and helix 3/10 (Figure 3.8) has been observed, while other residues reserve their original conformations during 6 ns MD runs. On the other hand, there are significantly conformational transitions observed during MD simulations for apo- α -lactalbumin, in particularly, at residues 17-20, 45-48 and 105-110 as seen in Figure 3.9. Residues 17-20 show transition from helix 3/10 into the turn conformations after 2 ns of MD runs, whereas residues 45-48 change their conformations between β sheet-like structures and the turn-like structures over the course of simulations. Another noticeably conformational alteration is observed in residues 105-110, which adopts from α helical conformation to the turn after 5 ns of MD runs. It is interesting that the residues from domains showing transitions in their structures during MD runs are located in the highly flexible regions, as shown in Figure 3.7. However, longer simulations may be needed to investigate structural changes especially in highly flexible regions of both apo- α -lactalbumin and lysozyme.

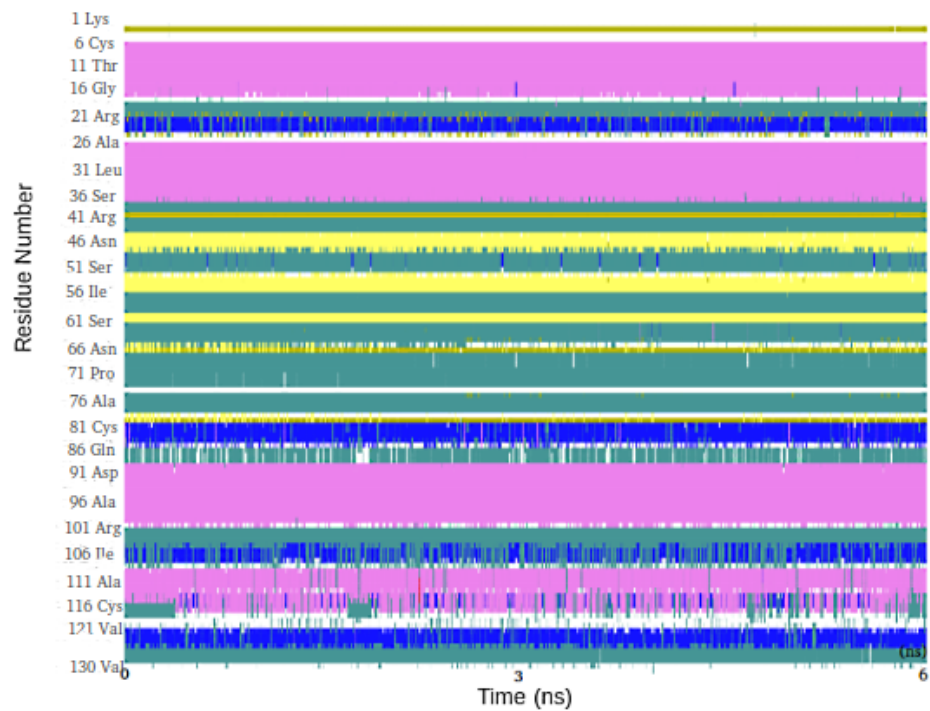


Figure 3.8: Time evolution of secondary structure of lysozyme, purple, pink, green, yellow, gold and white represent α helices, helices 3/10, turns, main- β sheets, partial- β sheets and coils, respectively.

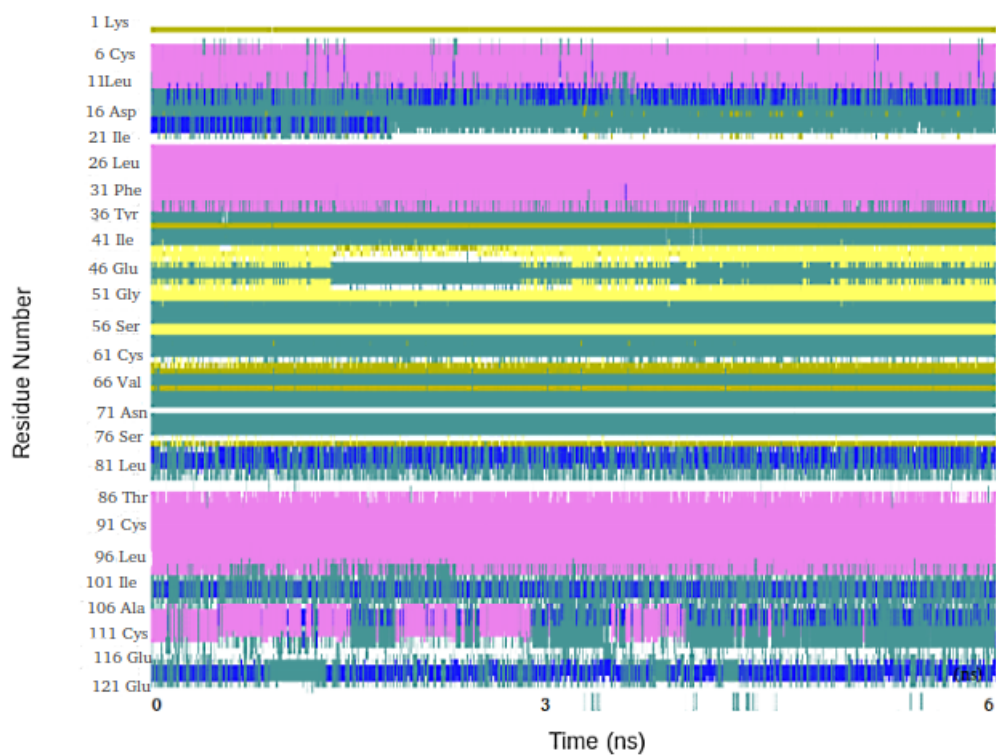


Figure 3.9: Time evolution of secondary structure of apo- α -lactalbumin, purple, pink, green, yellow, gold and white represent α helices, helices 3/10, turns, main- β sheets, partial - β sheets and coils, respectively.

Analysis of Radius of Gyration

Figure 3.10 shows evolution of the radius of gyration of lysozyme and apo- α -lactalbumin during 6 ns MD runs. Overall, the radius of gyration of both proteins does not show any significant changes over the simulation time. In the case of lysozyme, the radius of gyration starts at approximately 14 Å and slightly increases with around 0.1 Å for the first 3 ns. This is followed by a 0.1 Å decrease at 3.5 ns before gradually increasing to almost remain constant at approximately 14.1 Å during 4-6 ns MD runs. This indicates that lysozyme molecule is slightly expanded during 6 ns MD runs. In case of apo- α -lactalbumin, its radius of gyration decreases from around 14.1 Å to 13.8 Å after 3 ns MD runs before gradually increasing to 13.95 Å at 6 ns MD runs. Regarding to radius of gyration of apo- α -lactalbumin in Figure 3.10, the protein dimension is more compact with a 0.3 Å decrease in its dimension during the first 3 ns MD runs, then the protein molecule is expanded gradually with an around 0.15 Å increase in its dimension at the 6 ns MD runs.

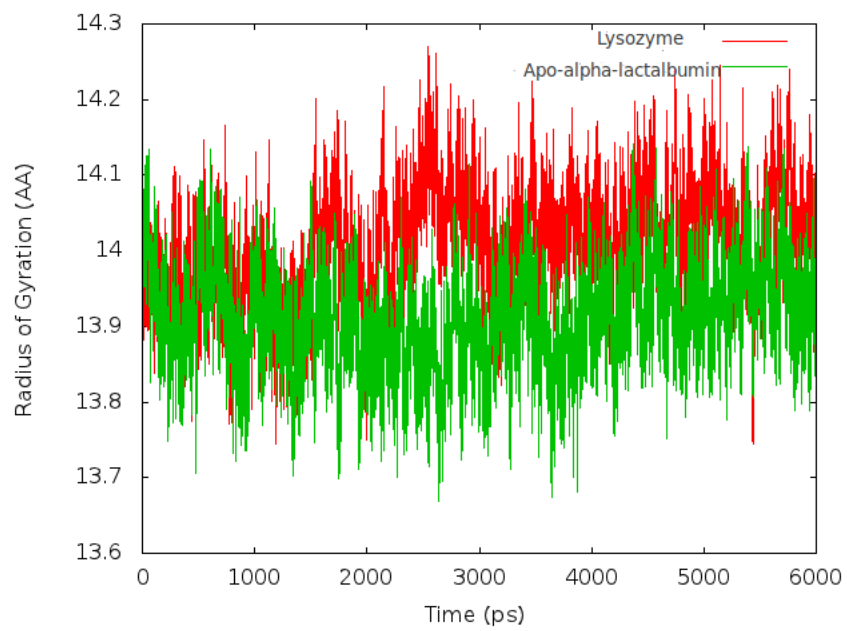


Figure 3.10: Radius of gyration of lysozyme and apo- α -lactalbumin.

3.4 Summary

Lysozyme and α -lactalbumin are proteins having comparable folding patterns with larger β -sheet than α -helical domains, shape, size and number of disulfide bonds formed in the protein molecules.³⁵ However, lysozyme poses higher structural stability than α -lactalbumin.^{28,35} X-Ray crystal structures of human lysozyme and human α -lactalbumin have been analyzed by using the VMD software. Solvent exposed residues, turns, coils and helix 3/10 located closer to the C-terminal show higher fluctuations compared to the other locations for both proteins. All-atom explicit solvent MD simulations have been carried out to analyze structural stability of lysozyme and apo- α -lactalbumin under physiological conditions (pH 7.0). Even though the backbone RMSD is not a useful tool to characterize stability of both proteins because the backbone RMSD of both proteins relative to their crystal structures and to their minimized structure showed different trends. Larger backbone RMSD of lysozyme than apo- α -lactalbumin relative to their crystallography structures was observed. On the other hand, when the backbone RMSDs of both proteins were calculated with respect to their minimized structures, larger backbone RMSD of apo- α -lactalbumin was noticed instead. The backbone RMSF does show higher fluctuations in the exposed loops and the C-terminal domain of apo- α -lactalbumin compared to lysozyme. RMSFs of residues with α helices and β sheets conformations are not significantly different. The evolution of the secondary structure of apo- α -lactalbumin also indicates that the protein partly denatures, especially at highly fluctuating residues 105-109. An analysis of radius of gyration of both proteins does not show a significant change in protein dimensions; hence, longer simulations may be suggested to observe this. In conclusion, the lysozyme molecule has higher stability compared to apo- α -lactalbumin in physiological environment, as it follows from all-atom explicit solvent simulations.

Chapter 4

Molecular Mechanics Simulations of Lysozyme and Apo- α -Lactalbumin Adsorption at Hydrophobic and Hydrophilic Surfaces

4.1 Introduction

Given the normalization of both human lysozyme and human α -lactalbumin from their crystal state to physiologically relevant solutions, it is now possible to further explore how they may interact with polymer surfaces. However, to perform atomistic simulations to mimic adsorption of a whole protein in real situation is very challenging because a protein is a complex bio-molecule composing of several thousand(s) atoms. Obviously, to investigate behaviours of all protein atoms is not a simple task. Moreover, adsorption of a protein at a given surface in real situation and in atomistic simulations are quite different. Since, in real situation the adsorption process only take few seconds after a biomaterial is immersed

into biofluids. These few seconds are processing time for a protein to diffuse to the surface and orient itself before adsorbing at the surface with its most preferable orientation. Nonetheless, in atomistic simulations a protein may not be able to self-adjust to adsorb at a biomaterial surface with most stable orientation as this process is typically time-consuming in simulations especially in explicit solvation models and a course of simulations is also usually limited. Therefore, to perform atomistic simulations of a whole protein molecule, which is generally chemically inhomogeneous, selection of the most preferable adsorption orientation at a given surface is commonly determined first, prior to any investigation of surface-induced adsorption or conformational changes of the protein. It must be noted that these two processes can be conducted separately because it has been shown that the rate of internal changes in conformations of a protein is usually faster than rate of surface-induced conformational changes within the protein.⁴⁶

In addition, to perform atomistic simulations to investigate surface-induced conformational changes of a protein, it is necessary to put the protein at a position where is optimum for adsorption to optimize usage of computational resources. One valuable example was reported by Wei et al.⁴⁵ using the adsorption of lysozyme at the hydrophobic polyethylene surface (PE). According to the study, 300 ns were needed for the investigation of a complete adsorption process, while lysozyme with selected orientation was put only 9 Å above the PE surface in explicit TIP4P water box. As addressed in the study, 10 ns were needed for the protein to diffuse to the surface and 70 ns were the time for dehydration induced by hydrophobic interactions before the protein adsorbed at the surface with its long axis parallel to the surface. The rest of simulation time (220 ns) was taken for observing the surface-induced conformational changes of the protein. Therein, the most stable orientation of the protein model and the optimum distance for adsorption at the surface model will be conducted first in the current project before applying MD simulations to investigate surface-induced conformational change of the proteins.

Although, MD simulations, together with use of explicit solvation mod-

els, yield more accurate results for determination of the most preferable adsorption orientation of a protein at a given surface than MC and MM simulations, MD methods are not commonly used for this application because they are very computationally expensive. Generally, MC simulations with use of continuum medium and treatment of a protein and polymer as rigid bodies are conducted.^{3,47,50} Nonetheless, MC simulations even already applied with continuum solvation model are still computationally expensive.⁵³ Therefore, MM simulations with continuum solvation model have been used by several authors instead.^{51,52,53,54,55} Moreover, the results from these authors also showed insignificant differences from those reported by conducting MC simulations and experiments.^{51,52,53,54,55} As a result of these successful MM simulations for selection of the most preferable adsorption orientation of a protein to adsorb at a given surface, the MM simulations combined with continuum Generalized Born model will be conducted for the investigation of initial adsorption orientation of lysozyme and apo-human- α -lactalbumin at the surface models. Details of MM simulations will be mentioned in the next section followed by simply analyzing vdW interaction energy arising from interactions between a protein approaching to the surface with different orientations and hydrophobic (-OCH₃) and hydrophilic (-OH) surfaces, respectively. An explanation of initial adsorption mechanism of the protein models at hydrophobic (-OCH₃) and hydrophilic (-OH) model surfaces will be left for future work because there is still some uncertainty in our results, due to the lack of convergence of the EL interaction energy term averaged over the simulation time period.

4.2 Methodology

4.2.1 Description of Polymer, Protein Models and Simulations

Thiolated PEO chains with the two terminal groups of -OH and -OCH₃ were obtained from Dr. Nikolay Blinolv. The polymer chains were pre-

pared using the software Accelrys Material Studio. The thiolated PEO chains capped with -OH and -OCH₃ were then prepared on Au(111) lattice by translating sulfur atoms attached to ethylene monomers (MW 222) and the monomers capped with -OCH₃ on all possible binding sites of the plane (111) of Au lattice or on $a(\sqrt{3} \times \sqrt{3})R30_z^\circ$ layer using tleap implemented in AMBER11 package ($a = 2.88 \text{ \AA}$, $R30_z^\circ = \text{rotated } 30^\circ \text{ from the Z axis perpendicular to the surface}$).⁴⁴ Accordingly, the distance between each sulfur atom translated in x-y plane was $\sim 0.5 \text{ nm}$ (Figure 4.1 a).⁴⁴ The total number of PEO-OH and PEO-OCH₃ chains prepared over the Au (111) layer were 324 chains. It must be noted that as reported by Godawat et al.¹⁰, an insertion of Au surface could be neglected if packing density of polymer was highly dense (i.e. a single chain of thiolated SAM was replicated at all possible sites of the plane (111) of Au.), since water over the polymer layer could not feel Au in the study. Accordingly, Au surface was not inserted in the current project, and only a sulfur atom attached to PEO-OH/ PEO-OCH₃ was replicated over Au (111) layer as described above.

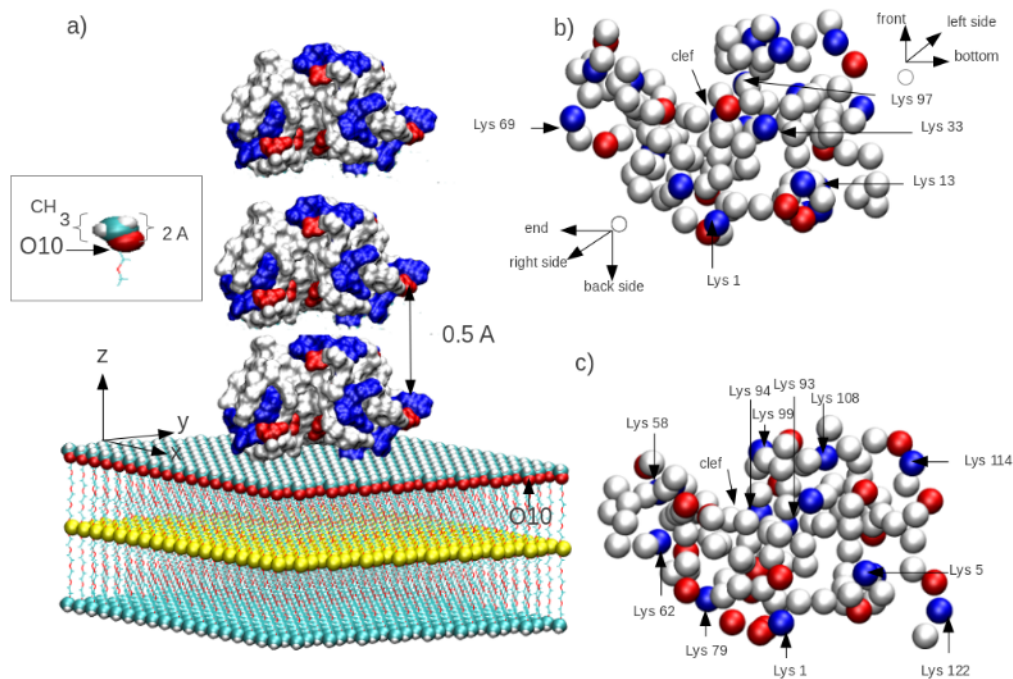
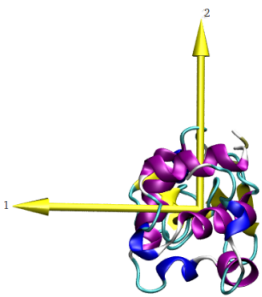
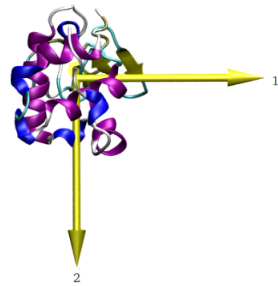


Figure 4.1: a) A complex system composed of a protein translated in z direction with an increment of 0.5 \AA and HS-PEO-OCH₃ translated in x - y plane depicted in the figure (there are three different things here: (i) protein translation as a part of MM protocol (ii) polymers translation as a part of the procedure used to build a surface (iii) chain length of OCH₃ group $\sim 2 \text{ \AA}$ attached to O10 atom). CH₃ groups, sulfur atoms, O10 atoms and PEO chains labeled as cyan, yellow, red vdW spheres and cyan lines, respectively. b) Graphical explanation of different parts of lysozyme molecule c) Graphical representation of different parts of apo- α -lactalbumin molecule.

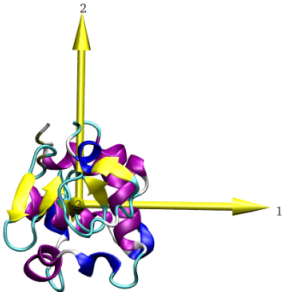
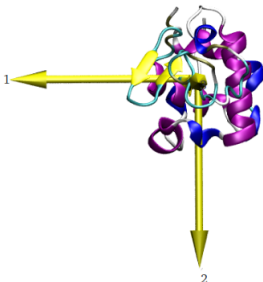
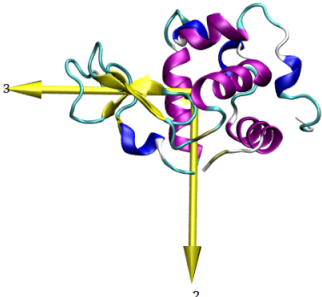
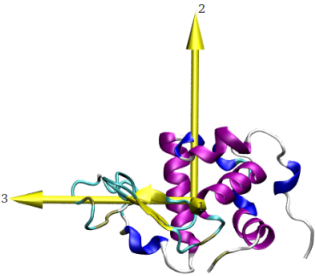
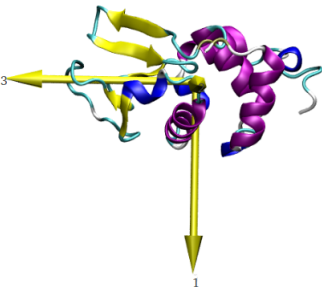
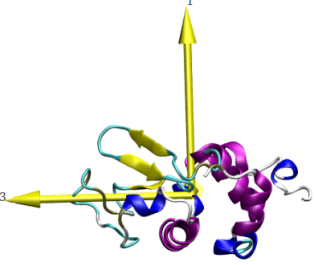
As discussed thoroughly in chapter 3, the initial crystal structures of the wild-type human lysozyme and human α -lactalbumin deposited in the Protein Data Bank with PDB ID of 1REX³³ and 1A4V³², respectively were used as initial crystal structures. pKa values of both proteins were then determined using the software PROKA 3.1 (see appendix A) to use as information for determining new charged states of the proteins before performing simulations at physiological pH with AMBER 11 package. In case of lysozyme, all acidic and basic residues were predicted to be deprotonated and protonated, respectively. Likewise, all acidic and basic residues of α -lactalbumin were also predicted to be deprotonated and protonated, respectively (see appendix A for pKa information). Subsequently, all water and ion molecules (including Ca⁺² ions found in α -lactalbumin molecule) were removed from the initial crystal structures of both proteins. Then, the principal axes of both proteins were set using external script provided in VMD website.⁷¹ Accordingly, crystal structures of both lysozyme and apo- α -lactalbumin with new coordinates were obtained. In the system of principal axes, the 1st and 2nd axes were set along the initial tensor of each protein. The 3rd axis was set in the direction of lowest moment of inertia of a protein (Table 4.1) and the origin point was chosen at center of mass of each protein. The starting orientation of lysozyme was then chosen for generating other lysozyme orientations and labeled as an initial structure as seen in Table 4.1. The lysozyme initial structure was then rotated clockwise in 90 ° increments around the 3rd axis with respect to its initial structure using the command (translate) implemented in antechamber package in AMBER 11. Consequently, four lysozyme orientations parallel to the long axis of the protein were generated and labeled as orientation 3, 4, 5 and 6 seen in in Table 4.1. The other two lysozyme orientations were generated by clockwise rotations of the protein about the 1st axis by 90 ° and 270 ° with respect to its initial structures. As a consequence, orientations 1 and 2 of lysozyme were generated (Table 4.1). In case of apo- α -lactalbumin, all protein orientations were chosen to superimpose all lysozyme orientations generated (Table 4.1). It must be noted that orientation 3, 4, 5 and 6 of both proteins were created based on results from

other simulations and experiments reported that a protein preferred to adsorb at a given surface with the orientation that areas of contact between the protein and the surface were maximize.^{45,46,53,57,66} On the other hand, lysozyme and apo- α -lactalbumin with orientation 1 and 2 were generated to test the hypothesis whether the proteins approaching to the surface with the orientations posing small areas of contact are less favorable to adsorb at a given surface than the proteins with orientations posing large areas of contact.

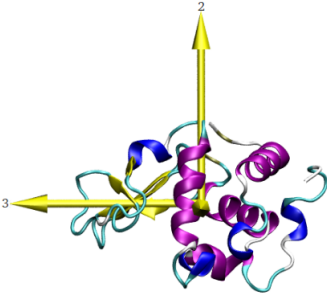
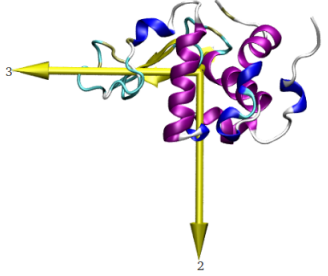
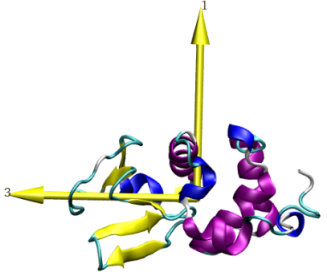
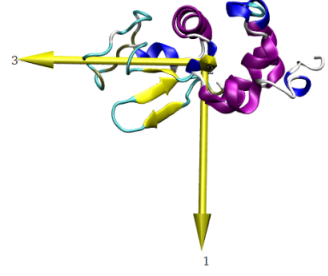
Table 4.1: Different orientations of lysozyme (left) and apo- α -lactalbumin (right). $R_3 90^\circ$, $R_3 180^\circ$, $R_3 270^\circ$, $R_3 360^\circ$ stand for 90° , 180° , 270° and 360° clockwise rotation about the 3rd axis, respectively. $R_1 90$ and $R_1 270$ stand for 90° and 270° clockwise rotation about the 1st axis, respectively. In cartoon representations, purple, blue, cyan, white, yellow and gold colours present α helix, helix 3/10, turn, coil, main $-\beta$ sheet and partial- β sheet conformations, respectively.

Lysozyme	Apo- α -lactalbumin
 <p data-bbox="402 1444 755 1480">1. [Initial Structure]$R_1 270^\circ$</p>	 <p data-bbox="873 1444 1226 1480">1. [Initial Structure]$R_2 270^\circ$</p>

continued on the next page

Lysozyme	Apo- α -lactalbumin
 <p data-bbox="412 699 743 730">2. [Initial Structure]$R_1 90^\circ$</p>	 <p data-bbox="883 699 1214 730">2. [Initial Structure]$R_1 90^\circ$</p>
 <p data-bbox="406 1176 753 1207">3. [Initial Structure]$R_3 180^\circ$</p>	 <p data-bbox="875 1176 1222 1207">3. [Initial Structure]$R_3 180^\circ$</p>
 <p data-bbox="406 1623 753 1654">4. [Initial Structure]$R_3 270^\circ$</p>	 <p data-bbox="875 1623 1222 1654">4. [Initial Structure]$R_3 270^\circ$</p>

continued on the next page

Lysozyme	Apo- α -lactalbumin
 <p data-bbox="454 724 706 756">5. Initial Structure</p>	 <p data-bbox="925 724 1177 756">5. Initial Structure</p>
 <p data-bbox="406 1186 747 1218">6. [Initial Structure]$R_1 90^\circ$</p>	 <p data-bbox="876 1186 1218 1218">6. [Initial Structure]$R_2 90^\circ$</p>

As expressed in Introduction section, if a protein was placed at sub-optimal positions for adsorption at a given surface even with its most preferable adsorption orientation, course of simulations would be still very lengthy since the protein would have to diffuse to the optimum position for the most stable adsorption. Therefore, the most preferable adsorption orientation and optimum position for adsorption of the protein models at the surface models were determined by first placing the proteins with different orientations at 10 \AA away from center of mass (ccm) of all sulfur atoms. Then each protein orientation was moved away from the ccm of the surface atoms with an increment of 0.5 \AA as simply depicted in Figure 4.1

a) until ccm of the protein and that of the sulfur atoms were separated at the distance of 30 Å. Accordingly, 480 complex systems were generated. It must be noted that the polymer surfaces were assumed to be homogeneous (translationally invariant in X-Y plane) ; therefore, additional degrees of freedom when the proteins with each orientation moved in x-y plane were neglected.

Next, the implicit solvation model was used in the simulations followed by conducting 500 MM steps with steepest descents method. It is noted that during the simulations, all surface atoms of the polymer were harmonically restrained using a force constant of 15 kcal/mol-Å². Subsequently, a single point free energy calculation without restraints applied was performed. Moreover, during simulations L-J interaction was set to be truncated at cut off value of 16.0 Å. Additionally, parameters for calculations of bond and non-bond interactions for the proteins and the polymer were taken from the ff99.r1 and GAFF force fields implemented in AMBER 11 package, respectively. However, atomic partial charges of the polymer were assigned externally by MOPAC quantum chemistry algorithm with bcc type as shown in Table A.1 and Table A.2, respectively.

It is worth to described graphical representation of lysozyme and apo- α - lactalbumin depicted in Figure 4.1b and c, respectively which will be useful for further explanation in this chapter. In Figure 4.1b), back side of lysozyme is defined as the face which Lys 1 is located being opposite to the front side where active site or clef is positioned. End side of the protein is defined as the face where Lys 69 is located which is opposite to the bottom of the protein (Figure 4.1b). Left and right sides of lysozyme are the faces where Lys 97 and Lys 33 are located, respectively. It must be noted that these graphical definitions of lysozyme have also been applied by several others.^{50,54,55} In case of apo- α - lactalbumin which has similar shape, size and around 35-40 % of amino acid sequences to lysozyme.³⁵ The definition of each part of lysozyme is also applied to apo- α -lactalbumin molecule as well. The faces of apo- α -lactalbumin where Lys 1 and clef are located are defined as back and front sides, respectively. End and bottom sides of the protein are defined as the faces where Lys 62 and Lys 112 are located,

respectively. Left and right sides of the protein are defined as the faces where Lys 95 and Lys 5 are located, respectively (Figure 4.1c).

4.2.2 Potential Energy

All-atom MM simulations were performed; however, some degrees of freedom (i.e. all surface atoms) were restrained. Also, change in free energy as a result of bonded interactions was small compared to non-bonded interactions. Therefore, only non-bonded terms expressed in Equation (2.4) in Chapter 2 are maintained for interaction energy calculations. EL interactions as described by Coulombs law (Chapter 2) were also modified for use with the continuum solvation model.⁶⁷ Overall potential energy was calculated using Equation (4.1):

$$U_{total} = \varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right) + \frac{e^2 q_i q_j e^{-\kappa r_{ij}}}{4\pi \epsilon_0 \epsilon_r(r_{ij}) r_{ij}} \quad (4.1)$$

In Equation 4.1, r_{ij} is the distance between atom i and j . ε_{ij} and σ_{ij} are L-J well depth and the distance between atoms i and j where the potential is minimum. e is electron charge, q_i and q_j are the partial charges on the atoms i and j , respectively. ϵ_0 is a relative permittivity of free space, while $\epsilon_r(r_{ij})$ is a distance-dependent relative permittivity which is assumed that $\epsilon_r(r_{ij}) = r_{ij}$ (in unit of Å) for distance less than 78.5 Å and = 78.5 for the distance greater than 78.5 Å. κ^{-1} is inverse Debye length and $\kappa^{-1} = \sqrt{\frac{\epsilon_0 \epsilon_r(r_{ij}) k_B T}{2 N_A e^2 I}}$, where k_B , N_A , T and I are Boltzmann constant, the Avogadro's number, temperature and ionic strength, respectively.⁵³ In all simulations, salt concentration is 0.2 M.

According to other simulations reported^{53,54,57}, dispersion interaction between a protein and uncharged or low charged density surface only acts over a range of ~ 10 -12 Å. Hence, the distance within 12 Å between ccm of the topmost atoms of a polymer surface (O10 and CH₃ for PEO-OH and PEO-OCH₃ surfaces, respectively.) and a protein/an amino acid atom was set as the most interacting region. In case of PEO-OH surface, the most interacting region was set within 12 Å from ccm of O10 atoms of PEO-OH. On the other hand, due to the chain length of CH₃ group attached to

O10 atom is $\sim 2 \text{ \AA}$ (see Figure 4.1); therefore, in case of PEO-OCH₃ surface the most interacting region was set within the distance of 14 \AA from ccm of O10 atoms. However, as mentioned in the last subsection that L-J truncation was set at 16 \AA . Hence, L-J interaction between protein atoms lying within 16 \AA from ccm of O10 atoms may be able to interact dispersively with O10 atoms of PEO-OH surface. Similarly, the protein atoms located within the distance of 18 \AA from ccm of O10 atoms may also be able to interact with CH₃ groups of PEO-OCH₃ surfaces. Additionally, the distance where the strongest dispersion interaction is observed is called a minimum distance following the definition applied by Hsu et. al.⁵³ while investigating L-J interaction between human serum albumin and hydrophilic-electrically neutral surface.

4.3 Results and Discussion

In this section, the most preferable adsorption orientation of lysozyme and apo- α -lactalbumin at PEO-OH and PEO-OCH₃ surfaces are aimed to be determined based on global minima of potential energy calculated by Equation (4.1). It must be noted that, some authors⁵³ determined the most preferable adsorption orientation of a protein at a hydrophilic-electrostatically neutral surface based on L-J interactions only. However, the most preferable adsorption orientation of the model proteins at the model surfaces in the current project will be determined based on global minima of the interaction energy, which is a combination of both L-J and EL interactions since atomic partial charge of the protein models and the surface models are non-zero. That said, only L-J contribution will be described herein since EL interactions have not been finalized. Therefore, a conclusion for the most preferable adsorption orientation of both protein models at the surface models will be left for future analysis in future work.

4.3.1 Effect of a Protein Orientation on Adsorption of Lysozyme at Hydrophobic PEO-OCH₃ surface

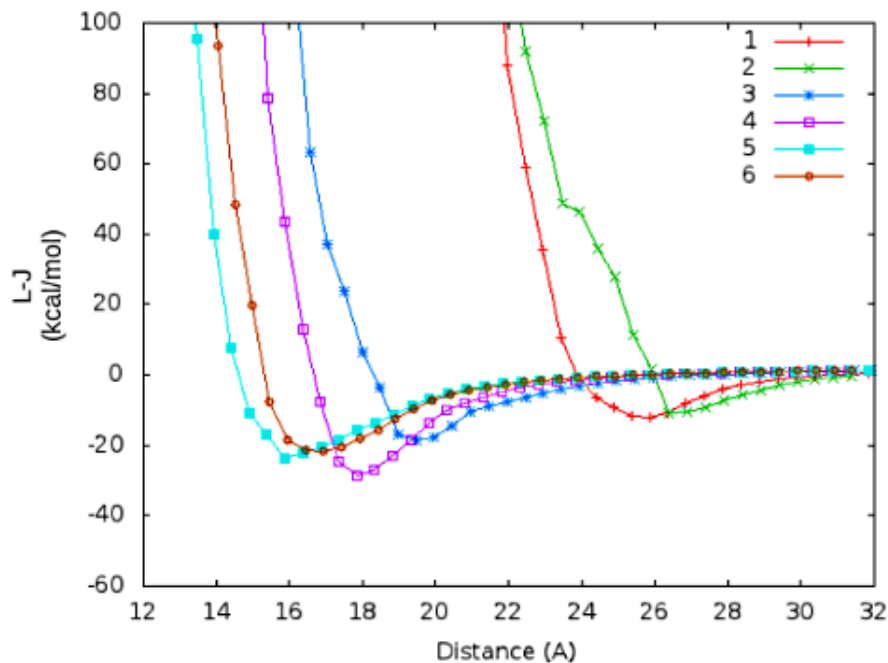


Figure 4.2: L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OCH₃ surface.

As mentioned in the last section, lysozyme with six orientations were positioned over PEO-OCH₃ surfaces at the distance varying from 10 to 30 Å from ccm of sulfur atoms. Total interaction energy which is a combination of both bonded and non-bonded interactions as described in Chapter 2 were calculated. However, the interaction energy arising from bonded interactions was small compared to non-bonded interactions (L-J and EL interactions). Therefore, only L-J and EL interactions were accounted for in total interaction energy calculation. Nonetheless, there is no convergence for EL interaction. Hence, only L-J energies arising from the interactions between lysozyme with different orientations and PEO-OCH₃ surface are shown in Figure 4.2. In Figure 4.2, as the protein approaches toward the surfaces, dispersion interaction increases and reaches the max-

imum value at a minimum distance specific for each protein with a specific orientation-surface interaction. Nevertheless, when the protein with a specific orientation approaches to the PEO-OCH₃ surface closer than its minimum distance, L-J energy increases dramatically due to overlapping of electron clouds of protein atoms and that of surface atoms. Regarding to Figure 4.2, the strongest dispersion energy arises from interaction between lysozyme with orientation 4 and PEO-OCH₃ with the energies of ~ -30 kcal/mol at a minimum distance of ~ 18 Å, while the least strongest dispersion energy arise from the interaction between the protein with orientation 2 and the surface with the energies of ~ -16 kcal/mol at a minimum distance of ~ 26.5 Å. Moreover, it is noticeable in Figure 4.2 that dispersion energies at minimum distances arise when lysozyme approaches to the surface with orientations which the long axis of the protein is parallel to the surface (i.e. 3,4, 5 and 6) are stronger (dispersion energies of ~ -20 to -30) kcal/mol) than that when lysozyme approaches to the surface with orientations 1 and 2 posing the long axis of the protein being perpendicular to the surface (dispersion energy of ~ -16 to -20 kcal/mol).

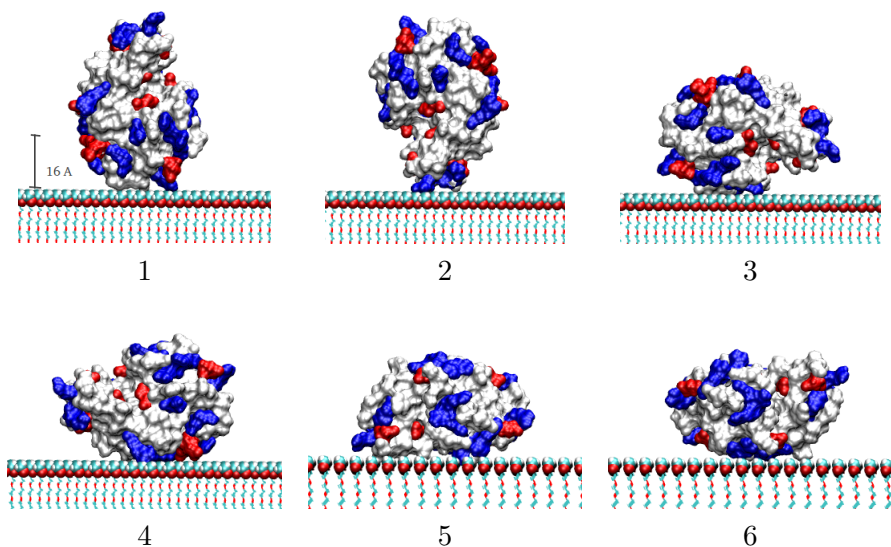


Figure 4.3: Different initial adsorption orientation of lysozyme at PEO-OCH₃ surface. Red, blue and white colours represent acidic, basic and other residues, respectively.

Strength of dispersion interaction was described by L-J attractive part in Equation (4.1). If the L-J parameters describing the interaction between each pair of atoms are not numerically markedly different, the strength of dispersion interaction should depend strongly on the inverse distance between each pair of atoms. Therefore, dispersion interaction between CH₃ groups of PEO-OCH₃ surface and the most exposed residues of lysozyme to the surface should be stronger than the interaction between the residues lying further away from the surface and the CH₃ groups. To analyze dispersion strength, the complex conformations which the strongest dispersion interaction between lysozyme with a specific orientation and the surface are observed are illustrated in Figure 4.3. Figure 4.3 shows conformations of lysozyme with six orientations over PEO-OCH₃ surface when ccm of lysozyme with orientations 1, 2, 3, 4, 5 and 6 are located at the minimum distances of 24.4, 26.4, 19.5, 17.8, 17 and 15.8 Å, respectively. These minimum distances correspond to L-J energy of $\sim -18, -16, -30, -$

25, -22 and -20 kcal/mol, respectively. In Figure 4.3, red, blue and white colours in lysozyme molecule represent acidic, basic and other residues, respectively which are depicted in the Figure for further analysis of EL interaction. Moreover, the distance between ccm of O10 atoms of PEO-OCH₃ and that of C_α of each lysozyme residue are also determined as seen in Figure 4.4a. In Figure 4.4a, ccm of CH₃ groups attached to O10 atoms is indicated at the distance of 2 Å over ccm of O10 atoms. Truncation for L-J interaction between CH₃ groups and lysozyme atoms is roughly indicated at the distance of 18 Å over ccm of O10 atoms. The most interacting region between CH₃ groups and lysozyme atoms is set within the distance of 14 Å over ccm of O10 atoms. Additionally, the total number of residues lying in the most interacting region is also given and divided into hydrophobic/hydrophilic neutral, acidic and basic groups in Figure 4.4b for further analysis when EL interaction and solvation effect will be accounted for in the total interaction energy calculation in future work.

Regarding to Figure 4.2, the strongest dispersion interaction with energy of ~ -30 kcal/mol arises from the interaction between the surface and lysozyme approaching to the surface with orientation 4. This large amount of energy may arise from the interaction between the lysozyme residues lying within the most interacting region and the surface. This is because lysozyme approaching to the surface with this orientation poses the highest number of residues lying in this region as summarized in Figure 4.4b. Moreover, the total number of residues lying in the most interacting region (shown as the numbers in parentheses) of lysozyme with orientation 4 (21 residues) is also higher than that of lysozyme with orientation 5 (19 residues), 6 (18 residues), and 3 (15 residues), respectively. This also corresponds to the rank of the dispersion energies at minimum distances which arise from the interactions between the surface and lysozyme approaching to the surface with orientation 4, 5, 6, and 3 with dispersion energies of ~ -30 , -25, -22 and -20 kcal/mol, respectively. However, it must be noted that the dispersion energy arises from the interaction between the surface and the other lysozyme residues lying further than the most interacting region but within L-J truncation region can also contribute to the total

dispersion energy as well. This can be clearly seen in case of dispersion interactions arise from the interactions between the surface and lysozyme approaching to the surface with orientation 1 and 2. Lysozyme approaching to the surface with orientation 1 poses smaller number of residues lying in the most interacting region (12 residues) than Lysozyme with orientation 2 (14 residues) as seen in Figure 4.4b . However, the dispersion energy at minimum distance arises from the interaction between the surface and lysozyme with orientation 1 (~ -18 kcal/mol) is stronger than that from the interaction between the surface and lysozyme with orientation 2 (~ -16 kcal/mol). This may be due to contribution to dispersion energy from the interaction between the surface and other lysozyme residues lying further than the most interacting region but within L-J truncation region. In Figure 4.4a, it is also noticeable that most residues of lysozyme approaching to the surface with orientation 2 lying further away from L-J truncation region than that of lysozyme moving to the surface with orientation 1. Therefore, it is reasonable that the dispersion energy arises from the interaction between lysozyme with orientation 2 and the surface is less strong than the dispersion energy from the interaction between lysozyme approaching to the surface with orientation 1 and the surface.

In conclusion, based on L-J energy lysozyme approaching to the surface with orientation 4 is the most preferable orientation for adsorption at hydrophobic PEO-OCH₃ surface. This lysozyme orientation is often called back on^{50,54} orientation which back side of the protein is positioned on the surface and cleft of the protein faces outward. The most interacting residues are residues 1-5 and 87-90 corresponding to coil and $\alpha 3$ conformations, respectively.

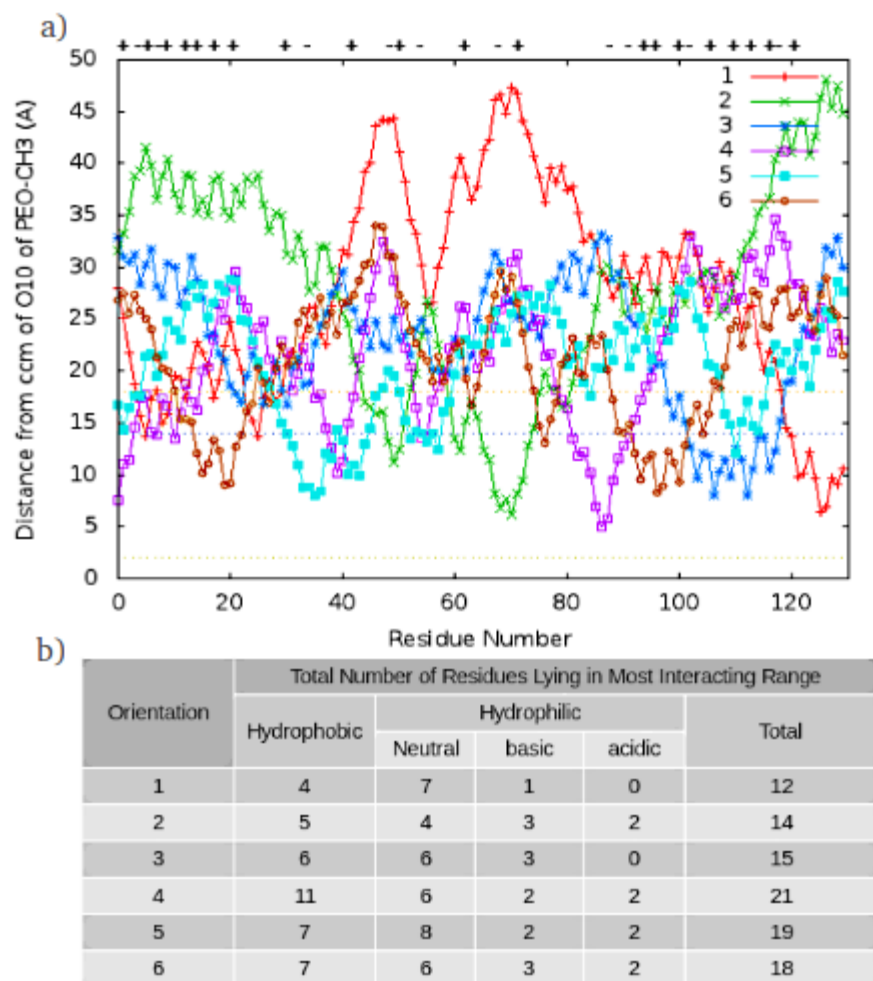


Figure 4.4: a) Residue distribution of each lysozyme orientation over PEO-OCH₃ surface including indications of the most interacting region (14 Å) and approximate L-J cut off region (18 Å) as well as ccm of CH₃ groups (2 Å). b) Summarization of total number of residues lying in the most interacting region.

4.3.2 Effect of a Protein Orientation on Adsorption of Lysozyme at Hydrophilic PEO-OH Surfaces

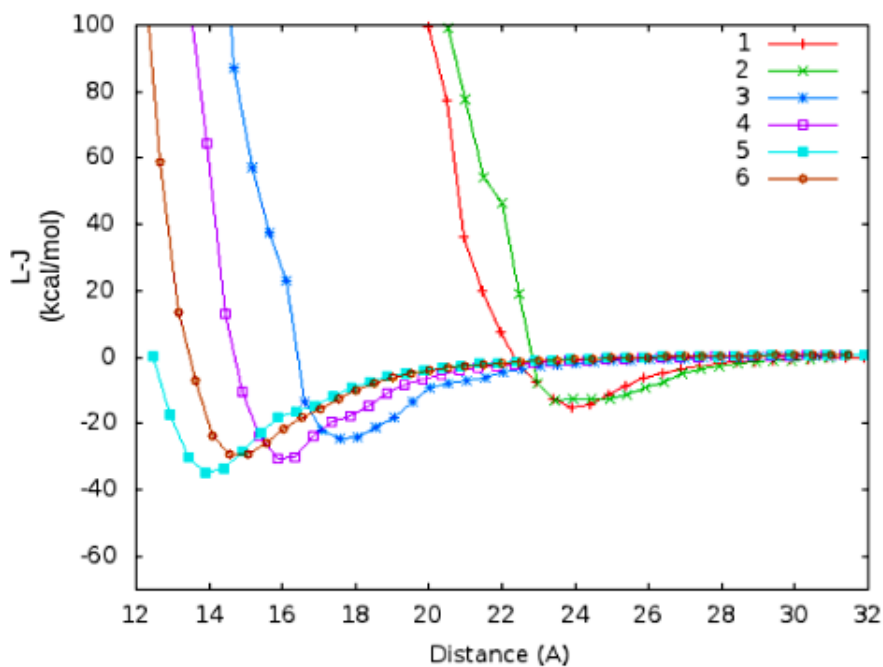


Figure 4.5: L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OH surface.

L-J interactions between the PEO-OH surface and lysozyme approaching to surface with different orientations calculated using the L-J potential well are shown in Figure 4.5. According to Figure 4.5, it can be seen that as the protein moves toward the surface, dispersion interaction increases gradually and reaches the maximum values at a minimum distance specific for each protein with a specific orientation-surface interaction. Nevertheless, when the protein with a specific orientation approaches to the surface closer than its minimum distance, L-J energy increases dramatically due to overlapping of electron clouds of protein atoms and that of surface atoms. Moreover, in Figure 4.5, it is noticed that the strongest dispersion inter-

actions of ~ -38 kcal/mol at minimum distance of ~ 14 Å from ccm of O10 atoms arises from the interaction between the PEO-OH surface and lysozyme approaching to the surface with orientation 5. Additionally, the least strongest dispersion interaction of ~ -18 kcal/mol at minimum distance of 24 Å from ccm of O10 atoms are observed from the interaction between the PEO-OH surface and lysozyme approaching to the surface with orientation 5 in the figure. Moreover, it is also noticeable in Figure 4.5 that the strength of dispersion interactions at minimum distances arising from interaction between the PEO-OH surface and lysozyme approaching to the surface with its long axis being parallel to the surface (orientations 3, 4, 5 and 6) are stronger (dispersion energy of ~ -25 to -35 kcal/mol) than the interactions arising when the protein moving toward the surface with orientations posing its long axis being perpendicular to the surface (orientations 1 and 2, dispersion energy of ~ -20 kcal/mol).

Dispersion strength may be simply analyzed by investigating residue distribution as done in the last subsection. Dispersion interactions between the PEO-OH surface (especially the topmost surface atoms i.e. O10 atoms) and protein atoms lying within L-J truncation region (16 Å from ccm of O10) can contribute to the total dispersion energy. However, the interaction arising from the protein residues lying in the most interacting region (12 Å from ccm of O10 atoms) interacting with the surface, may mainly contribute to the total dispersion interaction. To investigate the strength of dispersion interaction, complex conformations which the protein approaching to the surface with orientation 1, 2, 3, 4, 5 and 6 located at the distances of ~ 26 , 24, 18, 16, 14 and 15 Å over ccm of O10 atoms, respectively are of interest and illustrated in Figure 4.6 with red, blue and white colours depicted in lysozyme molecule representing acidic, basic and other residues for further analysis of EL interaction. In Figure 4.6, while lysozyme with orientations 1, 2, 3, 4, 5, and 6 are positioned over PEO-OH surface, it is roughly seen that there are more residues lying within L-J truncation region (16 Å) of lysozyme with orientation 5 than that of lysozyme with orientation 4, 6, 3, 1 and 2, respectively. Therefore, it is quite reasonable that dispersion interaction at minimum distance arises

from the interaction between the PEO-OH surface and lysozyme approaching to the surface with orientation 5 (-380 kcal/mol) is stronger than the interactions between the surface and lysozyme with orientations 4 (-35 kcal/mol), 6 (-30 kcal/mol), 3 (-28 kcal/mol), 1 (-20 kcal/mol) and 2 (-19 kcal/mol), respectively. Additional information of each lysozyme residue distribution over PEO-OH surface is shown in Figure 4.7 a. Summarization of the total number of residues lying in the most interacting region (16 Å over ccm of O10 atoms) for each lysozyme orientation is also given in Figure 4.7 b in different categories (i.e. hydrophobic, hydrophilic neutral, acidic and basic residues) for further analysis. Regarding to Figure 4.7 b, even though lysozyme approaching to the surface with orientation 6 poses the highest number of residues lying in the most interacting region, the strongest dispersion interaction at minimum distance arising from the interaction between the PEO-OH surface and lysozyme with orientation 5 is observed, instead. This can be explained that dispersion interaction between the surface and other lysozyme residues lying further than the most interacting region (12 Å from ccm of O10 atoms) but within the L-J truncation region (16 Å from ccm of O10 atoms) can also give rise to the total dispersion interaction.

In conclusion, based on L-J interaction lysozyme with orientation 5 is the most preferable adsorption orientation at PEO-OH surface. This lysozyme orientation is also called (right) side on orientation^{54,55} which its active site (clef) is parallel to the surface (Figure 4.6) The most interacting residues correspond to this lysozyme orientation are residues 30-45 and 53-58 relative to $\alpha 2$, turns, partial domain of $\beta 2$ and partial domain of $\beta 3$ conformations.

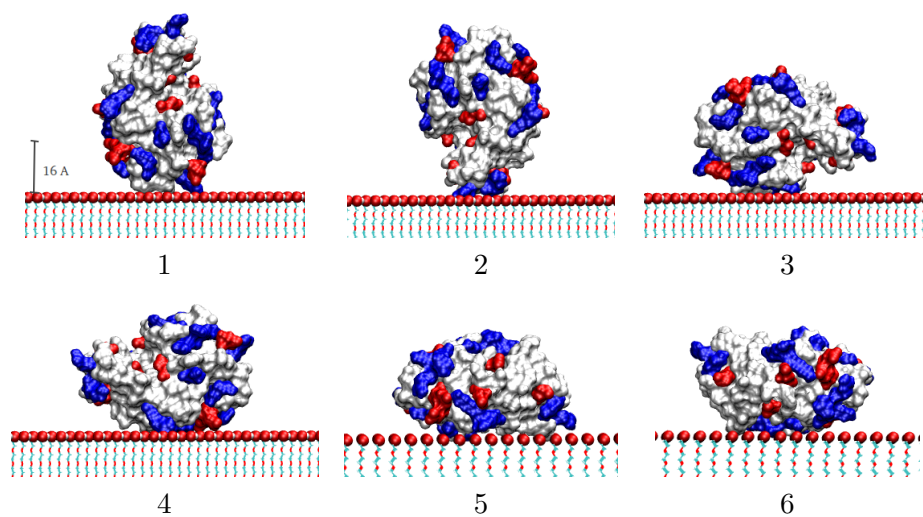


Figure 4.6: Different initial adsorption orientation of lysozyme at PEO-OH surface. Red, blue and white colours represent acidic, basic and other residues, respectively.

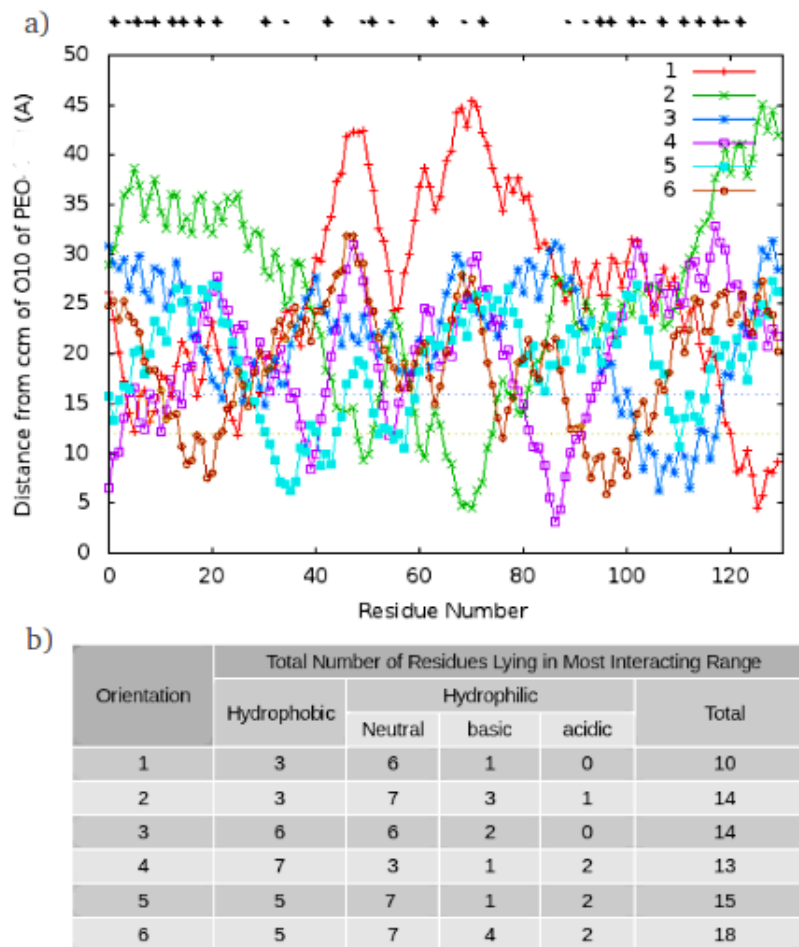


Figure 4.7: a) Residue distribution of each lysozyme orientation over PEO-OH surface with indications of approximate L-J interacting region and the most interacting region at the distance of 16 and 12 Å, respectively. b) Summarization of residues lying in most interacting region.

4.3.3 Effect of Orientation on Adsorption of Apo- α -lactalbumin at Hydrophobic PEO-OCH₃ Surface

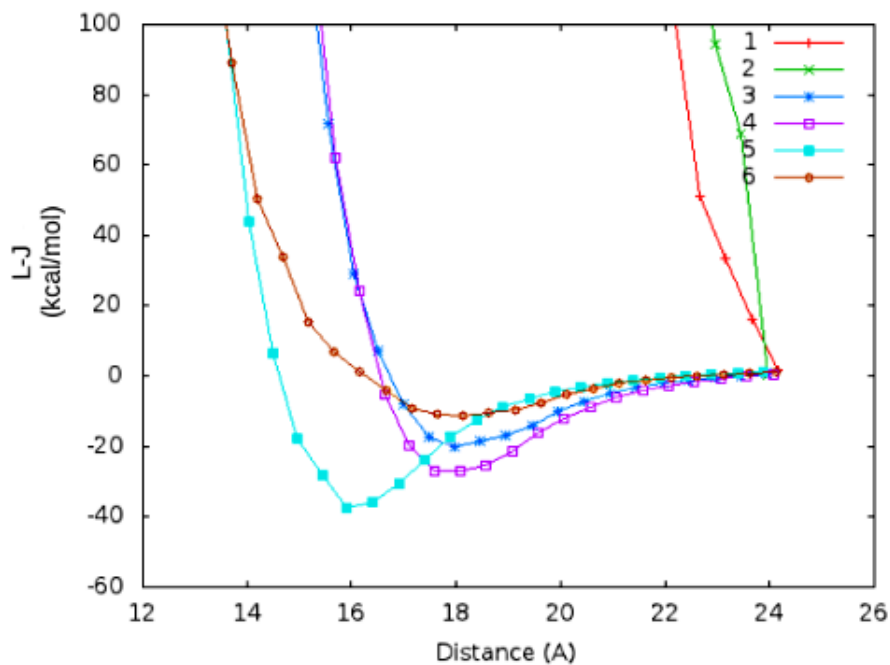


Figure 4.8: L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OCH₃ surface.

L-J interactions between the PEO-OCH₃ surface and apo- α -lactalbumin positioned over the surface with different orientations are shown in Figure 4.8. According to Figure 4.8, as the protein with orientations 3, 4, 5 and 6 approach to the surface, the dispersion interactions increase and reach the maximum values of ~ -20 , -30 , -38 , -16 kcal/mol at minimum distances of ~ 18 , 18 , 16 and 17.8 Å, respectively. If the protein with orientations 3, 4, 5 and 6 move closer to the surface than their minimum distances, strong repulsive forces arise due to overlapping of electron clouds of protein and surface atoms. In Figure 4.8, it is also noticed that total dispersion interaction cannot be observed when the protein with orientation 1 and 2

are positioned over ccm of O10 atoms of PEO-OCH₃ surface at the distances between ccm of O10 atoms and 24 Å (Figure 4.8). Accordingly, it is suggested that the reaction coordinate should be expanded for investigation of dispersion interactions between the PEO-OCH₃ surface and apo- α -lactalbumin approaching to the surface with orientation 1 and 2.

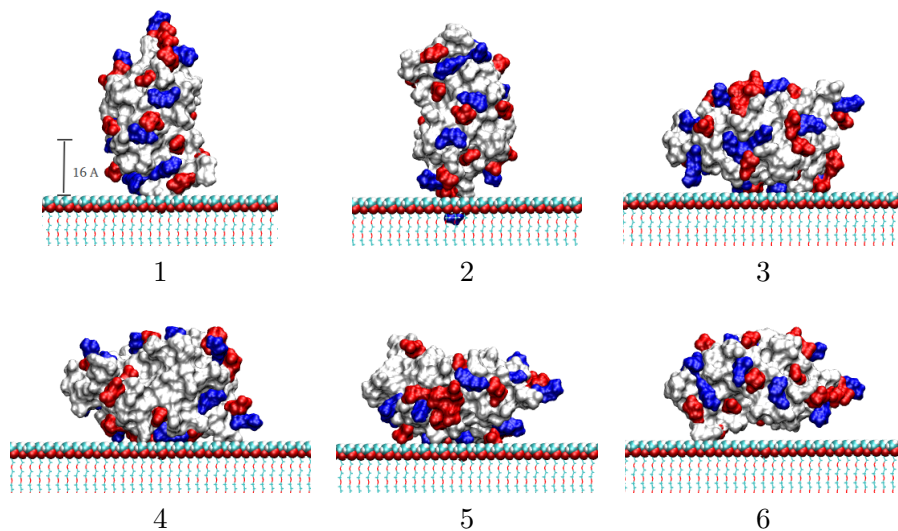


Figure 4.9: Different initial adsorption orientation of apo- α -lactalbumin at PEO-OCH₃ surface. Red, blue and white colours represent acidic, basic and other residues, respectively.

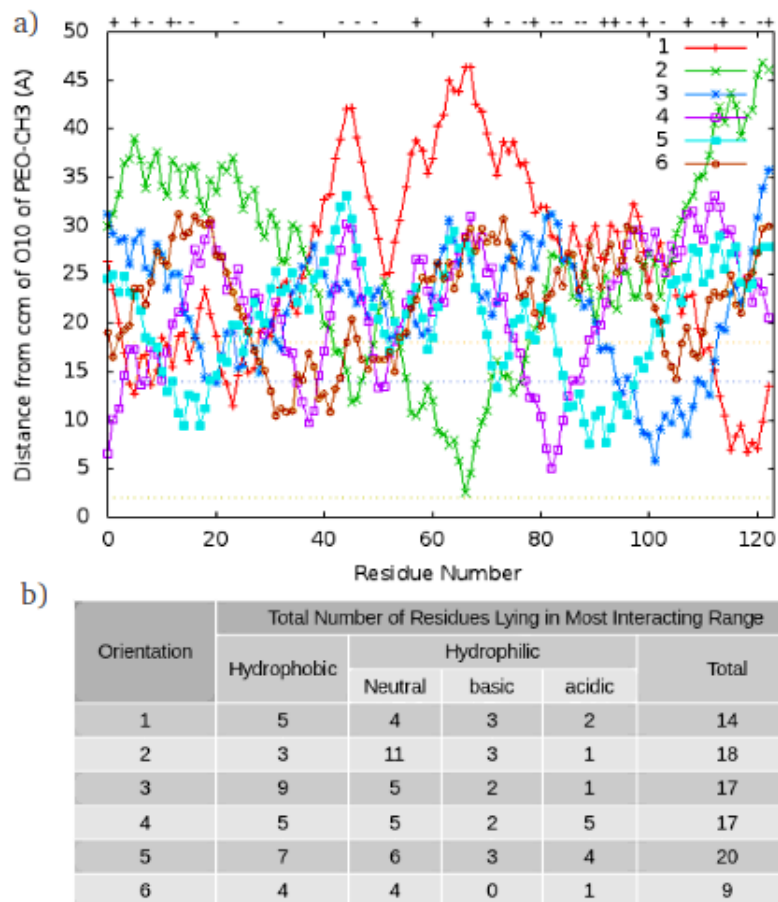


Figure 4.10: a) Residue distribution of each apo- α -lactalbumin orientation over PEO-OCH₃ surface including indications of the most interacting region (14 Å) and approximate L-J cut off region (18 Å) as well as ccm of CH₃ groups (2 Å). b) Summarization of total number of residues lying in the most interacting region.

In order to analyze dispersion strength, the complex conformations where dispersion interactions between the surface and the protein with different orientations are strongest are of interest and illustrated in Figure 4.8. In case of the protein approaching to the surface with orientation 1 and 2, total dispersion interactions between the protein and the surface cannot be observed (Figure 4.8). The complex conformations which the protein with orientation 1 and 2 exposed to the surface are positioned at the distances of $\sim 24 \text{ \AA}$ over ccm of O10 atoms are shown in Figure 4.9, instead. According to Figure 4.9, it is strongly reasonable that total dispersion interaction arise from the interaction between the surface and the protein approaching to the surface with orientations 2 cannot be observed. This is because some residues of the protein approaching to surface with orientation 2 are embedded into polymer layer (Figure 4.9). This leads to strong repulsive interaction which can contribute to the total L-J energy. Hence, only steric L-J energy is observed in this case. Regarding to Figure 4.9, it is also reasonable that dispersion interaction of $\sim -48 \text{ kcal/mol}$ arising from the interaction between the PEO-OCH₃ surface and the protein approaching to the surface with orientation 5 is stronger than the dispersion interactions arising from the interactions between the surface and the protein with orientation 4, 3 and 6, respectively. This is because it seems the protein with orientation 5 poses higher number of residues lying within L-J truncation region (18 \AA over ccm of O10) than the protein with orientation 4, 3, and 6, respectively (Figure 4.9). More details of residue distribution of apo- α -lactalbumin with different exposed orientations to the surface is given in Figure 4.10a. In Figure 4.10a, it is clear that most residues of the protein with orientation 1 and 2 positioned over the surface as shown in Figure 4.9 lie further away from the L-J truncation region (18 \AA over ccm of O10 atoms). Hence, it is reasonable that total dispersion interaction cannot be observed in these two cases. As apo- α -lactalbumin approaching to the surface with orientation 3, 4, 5 and 6 which the long axis of the protein being parallel to the surface (Figure 4.9), it can be seen that most protein residues lie in the L-J truncation region. Therefore, this also supports that strong dispersion interactions are noticed only in the

these cases.

As mentioned earlier that dispersion interaction arises from the interaction between the surface and the residues lying in the most interacting region (14 Å over ccm of O10 atoms) may also largely contribute to the total dispersion energy. Hence, the total number of residues lying in this region is summarized in Figure 4.10b. Details about hydrophobic, hydrophilic neutral, acidic and basic residues lying in this region are also given in Figure 4.10b for further analysis of EL interaction and solvation effect in future work. According to Figure 4.10b, the protein exposed to the surface with orientation 5 also poses highest number of residues (20 residues) lying in this region. Therefore, it is possible that the dispersion interaction between these residues and the surface may give large energy extent contributing to the total dispersion energy. However, the protein approaching to the surface with orientation 3 and 4 have the same number of residues lying in this region, dispersion interaction between the surface and the protein moving to the surface with orientation 4 (-30 kcal/mol) is stronger than that between the surface and the protein approaching to the surface with orientation 3 (-20 kcal/mol). This indicates that the interaction arises from the interaction between the surface and protein residues lying further than the most interacting region but within L-J truncation region can also give rise to the total dispersion energy.

In conclusion, based on L-J energy the most preferable adsorption orientation of apo- α -lactalbumin at PEO-OCH₃ surface is the protein with orientation 5 which is called side on orientation with clef being parallel to the surface and the left side of the protein lying over the surface. The most interacting residues are residues 10-20 and 85-100 (Figure 4.10a) corresponding to turn and α 3 conformations, respectively.

4.3.4 Effect of Orientation on Adsorption of Apo- α -lactalbumin at hydrophilic PEO-OH Surface

L-J interactions between the PEO-OH surface and apo- α -lactalbumin with different exposed orientations to the surface are illustrated in Figure 4.11. As the protein approaches to the surface with orientation 3, 4, 5 and 6

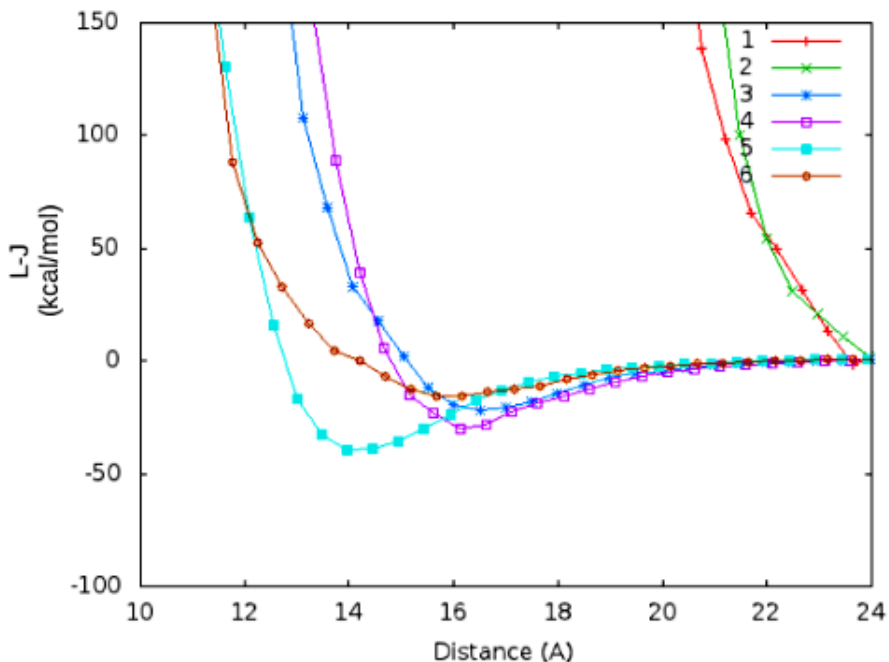


Figure 4.11: L-J energy as a function of distance measured from ccm of O10 atoms of PEO-OH surface

from the distance of around 24 Å, the dispersion interactions increase and reach the maximum values of -25, -30, -48, -20 kcal/mol at the minimum distances of ~ 16.8 , 16.4, 14 and 16 Å, respectively. It is also found in Figure 4.11 that total dispersion interactions between the surface and the protein approaching to the surface with orientation 1 and 2 cannot be observed over the reaction coordinates between ccm of O10 atoms and 24 Å. Hence, expansion of the reaction coordinates is suggested to observe the interaction at the distance further away from 24 Å.

Strength of dispersion interaction between the protein and the surface may be simply analyzed from residue distribution of the protein with different orientation lying over the PEO-OH surface but within L-J truncation region (16 Å from ccm of O10 atoms). Hence, the complex conformations of the surface and protein with orientation 3, 4, 5 and 6 located at the

minimum distances of 16.8, 16.4, 14 and 16 Å with energies of -25, -30, -48, -20 kcal/mol, respectively are investigated and illustrated in Figure 4.11. In case of L-J interaction between the surface and the protein with orientations 1 and 2 which total dispersion interactions cannot be noticed (Figure 4.11), the complex conformations when the protein is position at 24 Å are shown in Figure 4.12 instead. Regarding to Figure 4.12, it is reasonable that the dispersion interaction (-48 kcal/mol) at minimum distance (14 Å) between the surface and the protein approaching to the surface with orientation 5 is stronger than the interaction between the surface and the protein approaching to the surface with orientation 4 (-30 kcal/mol), 3 (-25 kcal/mol), and 6 (-20 kcal/mol), respectively. This is because the protein with orientation 5 seems to have more residues lying within L-J truncation region than protein with orientation 4, 3, and 6 as seen in Figure 4.12. Residue distribution of the protein approaching to the surface with different orientations is clearly shown in Figure 4.13a. Regarding to Figure 4.12, it is quite reasonable that dispersion interactions arise when the protein approaches to the surface with orientations 1 and 2 at the distance of 24 Å cannot be observed. This is because most residues of the protein with orientation 1 and 2 lying further than L-J truncation region. Moreover, residue 68 of the protein with orientation 2 is also embedded into the polymer layer (Figure 4.12). This may give large extent to the total L-J interaction leading to only steric L-J interaction is observed in this case (Figure 4.11).

As mentioned in the last subsection, the interaction between the surface and the protein residues lying within the most interacting region (12 Å over ccm of O10 atoms) may also contribute large energy extent to the total dispersion energy. The total number of protein residues lying in this region is summarized in Figure 4.13b. According to Figure 4.13b, the protein approaching to the surface with orientation 5 also poses the highest number of residues (19 residues) lying in this region. Therefore, it is possible that the dispersion interaction between these residues and the surface may give large energy extent contributing to the total dispersion energy. However, the protein approaching to the surface with orientation 3

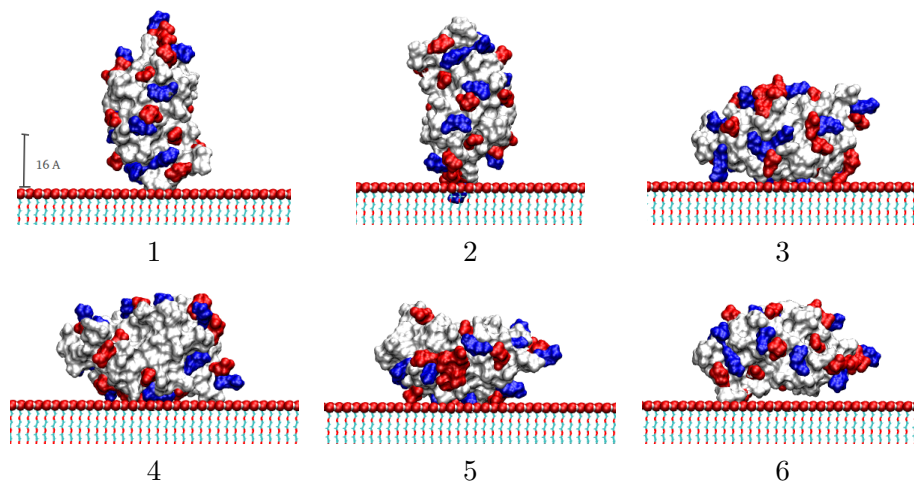


Figure 4.12: Different initial adsorption orientation of apo- α -lactalbumin at PEO-OH surface. Red, blue and white colours represent acidic, basic and other residues, respectively.

has higher number of residues lying in this region than the protein moving to the surface with orientation 4, its total dispersion (25 kcal/mol) energy is less strong than the later case (-30 kcal/mol). This indicates that the interaction between the surface and the residues lying further than the most interacting region but within L-J truncation region can also give rise to the total dispersion energy. It must be noted that the explanation of interaction between residues lying in the most interacting region divided into hydrophobic, hydrophilic neutral, acidic and basic residues in Figure 4.13b and the surface will be neglected herein. Since, this will be considered when EL interaction and solvation effect will be analyzed in future work.

In conclusion, based on L-J energy the most preferable adsorption orientation of apo- α -lactalbumin at PEO-OH surface is the protein with orientation 5 or side on orientation with clef being parallel to the surface and the left side of the protein exposed to the surface. The most interacting residues are residues 10-20 and 85-100 (Figure 4.13a) corresponding to turn and α 3 conformations, respectively.

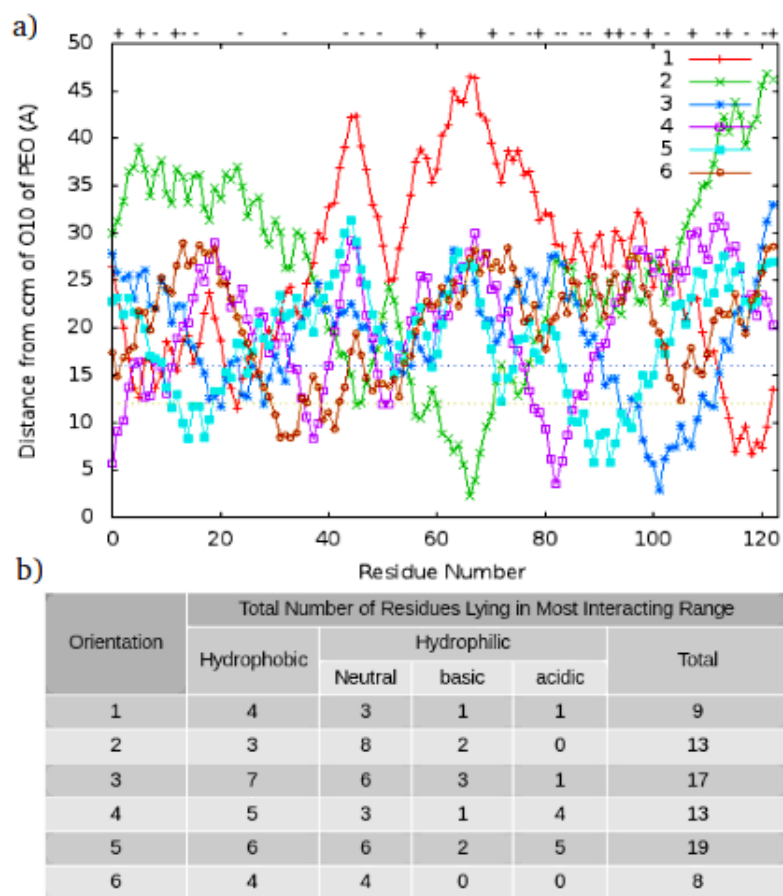


Figure 4.13: a) Residue distribution of each apo- α -lactalbumin orientation over PEO-OH surface with indication of approximate L-J and most interacting regions at the distance of 16 Å and 12 Å, respectively. b) Summarization of residues lying in most interacting region.

4.4 Summary

The aim of this work was to determine the preferred adsorption orientation of lysozyme and apo- α -lactalbumin at PEO-OH and PEO-OCH₃ surfaces using molecular mechanic (MM) simulations. This work is vital to the further investigation of adsorption behaviours of these two proteins, while in their most preferable adsorption orientations, at the surface modeled using molecular dynamic (MD) simulations. As mentioned, the most preferable adsorption orientation of the proteins at the surfaces will be determined based on a combination of L-J and EL interactions. EL interactions have not been converged yet; therefore, only L-J interactions have been analyzed in this chapter. EL interactions will be left for further analysis.

According to the calculations, it has been shown that the most preferable adsorption orientation of lysozyme at hydrophobic surface was back on orientation (cleft facing outward) and the most interacting residues corresponding to this protein orientation are residues 1-5 and 87-90 relative to coil and α 3 conformations, respectively. The most preferable adsorption orientation of lysozyme at hydrophilic surface was side on (clef being parallel to the surface) orientation which residues 30-45 and 53-58 relative to α 2, turn, partial β 2 and partial β 3 conformations are the most interacting residues and conformations, respectively.

In case of apo- α -lactalbumin, the most preferable adsorption orientation of the protein at hydrophobic and hydrophilic surfaces (based on L-J contributions only) were side on orientations (clef being parallel to the surface) which the most interacting residues are residues 10-20 and 85-100 corresponding to turn and α 3 conformations, respectively. The applications of MM simulations for determination of the most preferable adsorption orientation of a protein at a given surface could be further expanded for other applications such as identification of the most preferable binding site for antigen-antibody interaction in biosensor applications, drug delivery and enzyme-substrate interactions.

Chapter 5

Summary

5.1 Current Thesis

Molecular modelling is the best approach to theoretically describe some phenomena which cannot be reached in experimental scales including adsorption problems. Therefore, in the current study MM and MD simulations were conducted to investigate adsorption mechanisms of lysozyme and apo- α -lactalbumin which are considered as hard and soft protein, respectively at hydrophobic PEO-OCH₃ and hydrophilic PEO-OH surfaces. MD simulations were first performed separately to a system of lysozyme and apo- α -lactalbumin without a presence of polymer surface to observe changes in conformations of these 2 proteins in bulk solution. Regarding to these MD simulations, apo- α -lactalbumin showed lower internal stability than lysozyme with higher fluctuation of exposed residues and C-terminal including with highly noticeable conformational changes in residues 105-110 transitioning from α helices to turns observed in case of apo- α -lactalbumin but lysozyme during 6 ns of simulation time. Moreover, the results from these simulations also provided valuable information in that it is possible to investigate effect of surface-induced conformational changes of the proteins with a reasonable course of simulations.

Before proceeding to investigation of protein-surface interaction, selection of the most preferable initial adsorption orientation of the protein

models to the selected surface models by MM simulations were carried out. According to these simulations, all bonded and non-bonded interaction energy were calculated. However, the energy extent resulted from bonded interactions was small compared to that from non-bonded interactions. Therefore, only the energy arises from non-bonded interactions (L-J and EL interactions) were accounted for in total interaction energy for selection of the most preferable adsorption orientation of the protein models at the surface models based on global minimum energy. Nonetheless, EL interactions have not been converged yet; therefore, only L-J interactions have been analyzed. EL interactions will be left for further analysis. According to the calculations, the most preferable adsorption orientation of lysozyme at hydrophobic surface (based on L-J contributions only) was back on (cleft facing outward) orientation which the most interacting residues were residues 1-5 and 87-90 corresponding to coils and $\alpha 3$ conformations. At hydrophilic surface, lysozyme with side on (cleft being parallel to the surface) orientation which the most interacting residues were residues 30-45 and 53-58 relative to $\alpha 2$, turn, partial $\beta 2$ and partial $\beta 3$ conformation was the most preferable adsorption orientation based on L-J contributions only. In case of apo- α -lactalbumin, the most preferable adsorption orientation of the protein at hydrophobic and hydrophilic surfaces (based on L-J contributions only) were side on orientations (cleft being parallel to the surface) which the most interacting residues were residues 10-20 and 85-100 corresponding to turn and $\alpha 3$ conformations, respectively.

The applications of MM simulations for prediction of the most preferable adsorption orientation of a protein at a selected surface could be further expanded for other applications such as identification of the most preferable binding site for antigen-antibody interaction in biosensor applications, including drug delivery as well as enzyme-substrate interactions.

5.2 Future work

As a result, atomistic simulations applied to adsorption problems are still new areas compared to other problems such as folding-unfolding of a protein and drug delivery. There have been several difficulties must be solved during performing simulations. Moreover due to time for master student is limited. Therefore, further analysis of EL interactions from MM simulations is postponed to future work. Additionally, if convergence of EL interactions from MM simulations will be finalized, all energy extents arising from bonded and non-bonded interactions including solvation effect will be combined to total interaction energy. The adsorption mechanism of the protein models at hydrophobic and hydrophilic surfaces will be then investigated. Moreover, MD simulations for investigation of surface-induced conformational changes of the protein models will also conducted in future work.

Acknowledgements

Numerous discussion with Dr. Nikolay Blinolv is gratefully acknowledged. The research is partly supported by grants from the Royal Thai government.

Bibliography

1. Schmidt, D.R., Waldeck, H., Kao, W.J. (2009). Protein Adsorption to Biomaterials. In Puleo, D.A., Bizios, R. (Eds.). *Borrowed power: Biological Interactions on Materials Surfaces* (pp. 1-18). New York: Springer.
2. Leach, A.R. (2001). Energy Minimization and Related Methods for exploring the Energy Surface. *Borrowed power: Molecular Modelling: Principles and Applications* (pp. 253-301). Edinburgh: Pearson Education Limited.
3. Zheng, J., Li, L., Tsao, H-K., Sheng, Y-J., Jiang, J. (2005). A strong Repulsive Force between Protein and Oligo (Ethylene Glycol) Self-Assembled Monolayers: A Molecular Simulation Study. *Biophysical Journal*, 89(8), 159-166.
4. Norde, W. (1994). Protein Adsorption at Solid Surfaces: A Thermodynamic Approach. *Pure and Appl. Chem.*, 66(3), 491-494.
5. Ngadi, N., Abrahamson, J., Fee, C., Morison, K. (2009). Are PEG Molecules a Universal Protein Repellent?. *World Academy of Science, Engineering and Technology*, 49, 144-148.
6. Latour, R.A., Christopher J.R. (2002). Theoretical Analysis of Adsorption Thermodynamics for Hydrophobic Pep-

- tide Residues on SAM Surfaces of Varying Functionality. *J Biomed Mater Res*, 60(4), 564-577.
7. Rosenberg, K.J., Israelachvili, J. (2006). Recent Progress in Understanding Hydrophobic Interactions. *PNAS*, 103(43), 15739-15746.
 8. Galisteo, F., Norde, W. (1995). Adsorption of Lysozyme and α -Lactalbumin on Poly(styrenesulphonate) Lattices 1. Adsorption and Desorption Behaviours. *Colloids and surfaces B: Biointerfaces*, 4(6), 375-387.
 9. Kjellander, R., Florin, E. (1981). Water Structure and Changes in Thermal Stability of the System Poly(ethylene oxide)-Water. *J. Chem. Soc, Faraday Trans.1*, 77(9), 2053-5077.
 10. Godawat, R., Jamadagni, S.N., Garde, S. (2006). Characterizing Hydrophobicity of Interfaces by Using Cavity Formation, Solute Binding, and Water Correlations. *PNAS*, 106(36), 15119-15124.
 11. Bedrov, D., Smith, G.D. (2006). Molecular Dynamics Simulations Study of the Structure of Poly(ethylene oxide) Brushes on Nonpolar Surfaces in Aqueous Solution. *Langmuir*, 22(14), 6189-6194.
 12. Schroeder, D.V. (2000). Energy in Thermal Physics. *Borrowed power: An Introduction to Thermal Physics* (pp. 1-37). New York: Addison Wesley Longman.
 13. Hinchliffe, A. (2008). *Borrowed power: Molecular Modelling for Beginners*. Manchester: John Wiley and Sons Ltd.
 14. Allen, M.P., Tildesley, D.J. (1987). *Borrowed power: Computer Simulation of Liquids*. Oxford: Oxford University Press.

15. Sarupria, S., Garde, S. (2009). Quantifying Water Density Fluctuations and Compressibility of Hydration Shells of Hydrophobic Solutes and Proteins. *Phys. Rev. Lett.*, *103*(3), 37803-37806.
16. Davies, M., Williams, G., Loveduck, G.D. (1960). *Electrochem*, *64*, 575.
17. Unsworth, L.D., Sheardown, H., Brash, J.L. (2005). Protein Resistance of Surface Prepared by Sorption of End-Thiolated Poly(ethylene glycol) to Gold: Effect of Surface Chain Density. *Langmuir*, *21*(3), 1036-1041.
18. Frenkel, D., Smit, B. (2002). *Borrowed power: Understanding Molecular Simulation From Algorithms to Applications*. San Diego, California: Academic Press.
19. Dadarlat, V.M., Post, B.C. (2006). Decomposition of Protein Experimental Compressibility into Intrinsic and Hydration Shell Contributions. *Biophysical Journal*, *91*(12), 4544-4554.
20. Jamadagni, S.N., Godawat, R., Garde, S. (2009). How Surface Wettability Affects the Binding, Folding, and Dynamics of Hydrophobic Polymers at Interface. *Langmuir*, *25*(22), 13092-13099.
21. Lee, J.H., Lee, H.B., Andrade, J.D. (1995). Blood Compatibility of Polyethylene Oxide Surfaces. *Prog. Polym. Sci.*, *20*, 1043-1079.
22. Rein, P., Wolde, T., Chandler, D. (2002). Drying-induced Hydrophobic Polymer Collapse. *PNAS*, *99*(10), 6539-6543.
23. Basalyga, D.M., Latour, R.A. (2002). Theoretical Analysis of Adsorption Thermodynamics for Charged Peptide Residues on SAM Surfaces of Varying Functionality. *Borrowed power:*

Analysis of Adsorption Thermodynamics (pp. 120-130). Wiley Periodicals, Inc.

24. Raffianini, G., Ganazzoli, F. (2010). Protein Adsorption on a Hydrophobic Surface: A molecular Dynamics Study of Lysozyme on Graphite. *Langmuir*, 26(8), 5679-5689.
25. MacKerell, A.D, Jr, Bashford, D., Bellott, M., Dunbrack, R.L., Jr., Evanseck, J.D., Field, M.J., ... Fischer, S. (1998). All-Atom Empirical Potential for Molecular Modelling and Dynamics Studies of Proteins. *J. Phys.Chem B.*, 102(18), 3586-3616.
26. Latour, R.A. (2008). Molecular Simulation of Protein - Surface Interactions: benefits, problems, solutions and future directions. *Biointerphases*, 3(3), FC2-FC12.
27. Israelachvili, J. (1991). *Borrowed power: Intermolecular and Surface Force*. London: ACADEMIC PRESS.
28. Haynes, C.A., Norde, W. (1995). Structures and Stability of Adsorbed Proteins. *J. of Colloid and Interface Science*, 169(2), 313-328.
29. Roach, P., Farrar, D., Perry, C.C. (2005). Interpretation of Protein Adsorption: Surface-Induced Conformational Changes. *J. AM.CHEM.SOC*, 127(22), 8168-8173.
30. Hower, J.C., He, Y., Jiang, S. (2008). A Molecular Simulation Study of Methylated and Hydroxyl Sugar-Based Self-Assembled Monolayers Surface Hydration and Resistance to Protein Adsorption. *Journal of chemical physics*, 129(21), 215101-215107.
31. Xue, Y., Liu, J.N., Sun, Z., Ma, Z., Wu, C., Zhu, D. (2001). α -Lactalbumin Mutant Acting as Lysozyme. *PROTEINS: Structure, Function and Genetics*, 42(1), 17-22.

32. Chandra, N., Brew, K., Acharya, K.R. (1998). Structural Evidence for the Presence of a Secondary Calcium Binding Site in Human α -Lactalbumin. *Biochemistry*, *3*(14), 4767-4772.
33. Muraki, M., Harata, K., Sugita, N., Sato, K. (1996). Origin of Carbohydrate Recognition Specificity of Human Lysozyme Revealed by Affinity Labelling. *Biochemistry*, *35*(42), 13562-13567.
34. Horinek, D., Serr, A., Geisler, M., Pirzer, T., Slotta, U., Lud, S.Q., ... Garrido, J.A. (2008). Peptide Adsorption on a Hydrophobic Surface Results from an Interplay of Solvation, Surface and Intra Peptide Forces. *PNAS*, *105*(8), 2842-2847.
35. McKENZIE, H.A., White, JR. F.H. (1991). Lysozyme and α -lactalbumin: Structure, Function, and Interrelationships. *Adv. Protein Chem*, *41*, 173-315.
36. Privalov, P.L. (1979). Stability of Proteins: Small Globular Proteins. *Adv. Protein Chem*, *33*, 167-241.
37. Iler, K.L., Qasba, K.P. (1999). Molecular dynamic Simulations of α -Lactalbumin and Calcium Binding c-Type Lysozyme. *Protein engineering*, *12*(2), 129-139.
38. Zhou, R., Eleftheriou, M., Hon, C.C., Germain, R.S., Royyuru, A.K., Berne, B.J. (2008). Massively Parallel Molecular Dynamics Simulations of Lysozyme Unfolding. *IBM J.RES. and DEV*, *52*(1), 19-30.
39. Chrysina, E.D, Brew, K., Acharya, K.R. (2000). Crystal Structure of Apo-and Holo-bovine α -Lactalbumin at 2.2-Å Resolution Reveal an Effect of Calcium on Inter-lobe Interactions. *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, *275*(47), 37021-37029.

40. Kuehner, D.E., Engmann, J., Fergg, F., Wernick, M., Blanch, H.W., Prausnitz, J.M. (1999). Lysozyme Net Charge and Ion Binding in Concentrated Aqueous Electrolyte Solutions. *J. Physical Chem B*, 103(8), 1368-1374.
41. Smith, G.D., Bedrov, D., Borodin, O. (2000). Molecular Dynamic Simulations Study of Hydrogen Bonding in Aqueous Poly(Ethylene Oxide) Solutions. *PHYSICAL REVIEW LETTERS*, 85(26), 5583-5586.
42. Aray, Y., Marquez, M., Rodriguez, J., Vega, D., Simon-Manso, Y., Coll, S., ... Gonzalez, C. (2004). Electrostatics for Exploring the Nature of Hydrogen bonding in Polyethylene Oxide Hydration. *THE JOURNAL OF PHYSICAL CHEMISTRY B*, 108(7), 2418-2424.
43. Dormidontova, E.E. (2002). Role of Competitive PEO-Water and Water-Water Hydrogen Bonding in Aqueous Solution PEO Behavior. *Macromolecules*, 35(3), 987-1001.
44. Christopher, J.L., Estroff, L.A., Kriebel, J.K., Nuzzo, R.G., Whitesides, G.M. (2005). Self-Assembled Monolayers of Thiols on Metals as a Form of Nanotechnology. *Chem. Rev.*, 105(4), 1103-1169.
45. Wei, T., Carignano, M.A., Szleifer, I. (2011). Lysozyme Adsorption on Polyethylene Surfaces: Why Are Long Simulations Needed?. *Langmuir*, 27(19), 12074-12081.
46. Agashe, M., Vivek, R., Sturt, S.J., Latour, R.A. (2005). Molecular Simulation to Characterize the Adsorption Behavior of a Fibrinogen γ -Chain Fragment. *Langmuir*, 21(3), 1103-1117.
47. Hower, J.C., He, Y., Bernardis, M.T., Jiang, S. (2006). Understanding the Nonfouling Mechanism of Surfaces through

- Molecular Simulations of Sugar-based Self-assembled Monolayers. *The Journal of Chemical Physics*, *125*(21), 214704-214710.
48. Han, D.K., Park, K.D., Ahn, K.D., Jeong, S.Y., Kim, Y.H. (1989). Preparation and Surface Characterization of PEO-grafted and Heparin-Immobilized Polyurethanes. *J. Biomed. Mater. Res.: Applied biomaterials*, *23*(A1), 87-104.
 49. Jeon, S.I., Andrade, J. D. (1991). Protein-Surface Interactions in the Presence of Polyethylene Oxide. II Effect of Protein Size. *Journal of Colloid and Interface Science*, *142*(1), 159-166.
 50. Zhou, J., Chen, S., Jiang, S. (2003). Orientation of Adsorbed Antibodies on Charged Surfaces by Computer Simulation Based on a United-Residues Model. *Langmuir*, *19*(8), 3472-3478.
 51. Raffaini, G., Ganazzoli, F. (2009). Protein Adsorption on a Hydrophobic Surface: A Molecular Dynamics Study of Lysozyme on Graphite. *Langmuir*, *26*(8), 5679-5689.
 52. Raffaini, G., Ganazzoli, F. (2006). Protein Adsorption on the Hydrophilic Surface of a Glassy Polymer: a Computer Simulation Study. *Phys. Chem. Chem Phys*, *8*(23), 2765-2772.
 53. Hsu, H-J., Sheh-Yi, Tsay, R-Y. (2008). Preferred Orientation of Albumin Adsorption on a Hydrophilic Surface from Molecular Simulation. *Colloids and Surfaces B: Biointerfaces*, *67*(2), 183-191.
 54. Sun, Y., Welsh, W. J., Latour, R.A. (2005). Prediction of the Orientations of Adsorbed Protein Using an Empirical Energy Function with Implicit Solvation. *Langmuir*, *21*(12), 5616-5626.

55. Thyparambil, A.A. (2010). Structural Bioinformatic Based Method for Predicting the Initial Adsorbed Protein Orientation on a Surface (Master Thesis). Available from etd.lib.clemson.edu/.../Thyparambil_clemson_0050M_10775.pdf.
56. Raffaini, G., Ganazzoli, A. (2003). Simulation Study of the Interaction of Some Albumin Subdomains with a Flat Graphite Surface. *Langmuir*, 19(8), 3403-3412.
57. Noinville, V., Vidal-Madjar, C., Sebille, B. (1995). Modelling of Protein Adsorption on Polymer Surfaces. Computation of Adsorption Potential. *J. Phys. Chem.*, 99(5), 1516-1522.
58. Xie, Y., Zhou, J., Jiang, S. (2010). Parallel Tempering Monte Carlo Simulations of Lysozyme Orientation on Charged Surface. *The journal of Chemical Physics*, 132 (6), 65101-65108.
59. Protein Data Bank. (2012). *Electronic references*. Retrieved from <http://www.rcsb.org/pdb/home/home.do>.
60. Ryckaert, J.P., Ciccotti, G. (1977). Numerical Integration of the Cartesian Equation of Motion of a System with Constraints: Molecular Dynamics of nalkanes. *J. Comput. Phys*, 23, 327341.
61. Andersen, H.C. (1983). Rattle: A velocity version of the Shake algorithm for molecular dynamics calculations. *J. Comput. Phys*. 52, 2434.
62. Glattli, A., Daura, X. (2002). Derivation of an Improved Simple Point Charge Model for Liquid Water: SPC/A and SPC/L. *J. Chem. Phys*, 116, 98119828.
63. Mark, P, Nilsson, L. (2001). Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys.Chem. A*, 105, 99549960.

64. Jorgensen, W.L, Chandrasekhar, J. (1983). Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.*, 79, 926935.
65. Jorgensen, W.L, Madura, J.D. (1985). Temperature and Size Dependence for MonteCarlo Simulations of TIP4P Water. *Mol. Phys.*, 56, 13811392.
66. Dimer, F., Hubbuch, J. (2007). A Novel Approach to Characterize the Binding Orientation of Lysozyme on Ion-Exchange Resins. *Journal of Chromatography A*, 1149, 312320.
67. Schaefer, M., Bartels, C., Karplus, M. (1999). Solution Conformations of Structured Peptides: Continuum Electrostatics Versus Distance-Dependent Dielectric Functions. *Theor Chem Acc.* 101, 194-204.
68. Hjorteland, T. (1999). Method of Steepest Descents. *Electronic references*. Retrieved from <http://trond.hjorteland.com/thesis/node26.html>.
69. Hjorteland, T. (1999). Method of Conjugate Gradients. *Electronic references*. Retrieved from <http://trond.hjorteland.com/thesis/node27.html>.
70. Western Canada Research Grid. (2012). *Electronic references*. Retrieved from <http://www.westgrid.ca>.
71. NIH Resource for Macromolecular Modelling and Bioinformatics. (2012). *Electronic references*. Retrieved from <http://www.ks.uiuc.edu/Research/vmd>.

Appendix A

pKa of each Lysozyme residue determined by the PROKA 3.1 software

	Group	pKa	model-pKa		Group	pKa	model-pKa
	ASP 18 A	3.80	3.80		TYR 38 A	11.82	10.00
	ASP 49 A	2.47	3.80		TYR 45 A	10.11	10.00
	ASP 53 A	4.58	3.80		TYR 54 A	12.12	10.00
	ASP 67 A	2.24	3.80		TYR 63 A	10.18	10.00
	ASP 87 A	2.77	3.80		TYR 124 A	9.80	10.00
	ASP 91 A	3.61	3.80		LYS 1 A	11.29	10.50
	ASP 102 A	3.15	3.80		LYS 13 A	11.00	10.50
	ASP 120 A	2.53	3.80		LYS 33 A	10.06	10.50
	GLU 4 A	4.75	4.50		LYS 69 A	10.63	10.50
	GLU 7 A	3.01	4.50		LYS 97 A	10.00	10.50
	GLU 35 A	7.31	4.50		ARG 5 A	12.28	12.50
	C- 130 A	2.82	3.20		ARG 10 A	12.31	12.50
	HIS 78 A	6.48	6.50		ARG 14 A	12.19	12.50
	CYS 6 A	99.99	9.00		ARG 21 A	12.17	12.50
	CYS 30 A	99.99	9.00		ARG 41 A	12.39	12.50
	CYS 65 A	99.99	9.00		ARG 50 A	12.53	12.50
	CYS 77 A	99.99	9.00		ARG 62 A	12.59	12.50
	CYS 81 A	99.99	9.00		ARG 98 A	12.98	12.50
	CYS 95 A	99.99	9.00		ARG 101 A	13.30	12.50
	CYS 116 A	99.99	9.00		ARG 107 A	12.12	12.50
	CYS 128 A	99.99	9.00		ARG 113 A	12.20	12.50
	TYR 20 A	9.19	10.00		ARG 115 A	12.13	12.50
					ARG 119 A	12.33	12.50
					ARG 122 A	13.62	12.50
					N+ 1 A	7.90	8.00

pKa of each α -lactalbumin residue determined by the PROKA 3.1 software

Group	pKa	model-pKa	Group	pKa	model-pKa
ASP 14 A	3.83	3.80	CYS 6 A	99.99	9.00
ASP 16 A	3.18	3.80	CYS 28 A	99.99	9.00
ASP 37 A	3.23	3.80	CYS 61 A	99.99	9.00
ASP 74 A	3.85	3.80	CYS 73 A	99.99	9.00
ASP 78 A	2.84	3.80	CYS 77 A	99.99	9.00
ASP 82 A	3.17	3.80	CYS 91 A	99.99	9.00
ASP 83 A	2.90	3.80	CYS 111 A	99.99	9.00
ASP 84 A	2.41	3.80	CYS 120 A	99.99	9.00
ASP 87 A	3.50	3.80	TYR 18 A	11.43	10.00
ASP 88 A	7.25	3.80	TYR 36 A	11.86	10.00
ASP 97 A	3.61	3.80	TYR 50 A	10.22	10.00
ASP 102 A	2.83	3.80	TYR 103 A	13.30	10.00
GLU 7 A	3.13	4.50	LYS 1 A	11.18	10.50
GLU 25 A	3.79	4.50	LYS 5 A	10.21	10.50
GLU 43 A	4.65	4.50	LYS 13 A	10.54	10.50
GLU 46 A	4.65	4.50	LYS 58 A	10.36	10.50
GLU 49 A	5.72	4.50	LYS 62 A	10.41	10.50
GLU 113 A	4.57	4.50	LYS 79 A	10.70	10.50
GLU 116 A	4.42	4.50	LYS 93 A	10.30	10.50
GLU 121 A	4.49	4.50	LYS 94 A	10.26	10.50
C- 123 A	3.31	3.20	LYS 99 A	11.14	10.50
HIS 32 A	6.09	6.50	LYS 108 A	10.27	10.50
HIS 107 A	5.75	6.50	LYS 114 A	10.51	10.50
CYS 6 A	99.99	9.00	LYS 122 A	10.43	10.50
			ARG 70 A	12.41	12.50
			N+ 1 A	7.21	8.00

Table A.1: Atom name, type and atomic partial charge assigned by MOPAC (bcc) for PEO-OCH₃

Atom Name (Type)	Charge	Atom Name (Type)	Charge
H5 (h1)	0.0214	O6 (os)	-0.3572
H2B (h1)	0.0602	C7 (c3)	0.1101
H2A (h1)	0.0212	H18 (h1)	0.0372
C2 (c3)	0.1186	H9 (h1)	0.0676
O1 (os)	-0.4247	C15 (c3)	0.1277
H1A (h1)	0.0396	H33 (h1)	0.0373
H1B (h1)	0.0407	H26 (h1)	0.0624
C1 (c3)	0.1249	O7 (os)	-0.4295
H4 (h1)	0.0420	C8 (c3)	0.1281
H14 (h1)	0.0428	H19 (h1)	0.0480
C3 (c3)	0.1299	H10 (h1)	0.0385
O2 (os)	-0.4306	C16 (c3)	0.1301
H23 (h1)	0.0399	H34 (h1)	0.0453
H30 (h1)	0.0457	H27 (h1)	0.0396
C12 (c3)	0.1296	O8 (os)	-0.4299
H6 (h1)	0.0399	C9 (c3)	0.1297
H15 (h1)	0.0453	H20 (h1)	0.0453
C4 (c3)	0.1298	H11 (h1)	0.0399
O3 (os)	-0.4300	C17 (c3)	0.1297
H24 (h1)	0.0396	H35 (h1)	0.0457
H31 (h1)	0.0453	H28 (h1)	0.0399
C13 (c3)	0.1302	O9 (os)	-0.4307
H7 (h1)	0.0385	C10 (c3)	0.1300
H16 (h1)	0.0480	H21 (h1)	0.0429
C5 (c3)	0.1281	H12 (h1)	0.0419
O4 (os)	-0.4295	C18 (c3)	0.1248
H25 (h1)	0.0624	H36 (h1)	0.0408
H32 (h1)	0.0375	H29 (h1)	0.0395
C14 (c3)	0.1277	O10 (os)	-0.4247
H8 (h1)	0.0677	C11 (c3)	0.1186
H17 (h1)	0.0373	H22 (h1)	0.0212
C6 (c3)	0.1099	H13 (h1)	0.0602
O5 (os)	-0.3569	H3 (h1)	0.0215
S1 (ss)	0.2565		

Table A.2: Atom name, type and atomic partial charge assigned by MOPAC (bcc) for PEO-OH

Atom Name (Type)	Charge	Atom Name (Type)	Charge
H3 (ho)	0.3980	H31 (h1)	0.0442
O1 (oh)	-0.5928	H17 (h1)	0.0443
H1A (h1)	0.0778	H4 (ho)	0.3979
H1B (h1)	0.0417	O10 (oh)	-0.5929
C1 (c3)	0.1251	H11 (h1)	0.0779
H2B (h1)	0.0423	H18 (h1)	0.0415
H2A (h1)	0.0259	C9 (c3)	0.1251
C2 (c3)	0.0944	H25 (h1)	0.0423
O2 (os)	-0.4258	H32 (h1)	0.0257
H5 (h1)	0.0410	C16 (c3)	0.0944
H12 (h1)	0.0435	O9 (os)	-0.4258
C3 (c3)	0.1286	H10 (h1)	0.0401
H19 (h1)	0.0422	C8 (c3)	0.1286
H26 (h1)	0.0438	H24 (h1)	0.0417
C10 (c3)	0.1301	C15 (c3)	0.1300
O3 (os)	-0.4300	O8 (os)	-0.4299
H6 (h1)	0.0443	H9 (h1)	0.0404
H13 (h1)	0.0415	H16 (h1)	0.0456
C4 (c3)	0.1300	C7 (c3)	0.1296
H20 (h1)	0.0452	H23 (h1)	0.0407
H27 (h1)	0.0427	H30 (h1)	0.0467
C11 (c3)	0.1292	C14 (c3)	0.1290
O4 (os)	-0.4278	O7 (os)	-0.4295
H7 (h1)	0.0441	H8 (h1)	0.0525
H14 (h1)	0.0478	H15 (h1)	0.0423
C5 (c3)	0.0952	C6 (c3)	0.1282
H21 (h1)	0.0589	H22 (h1)	0.0625
H28 (h1)	0.0881	H29 (h1)	0.0588
C12 (c3)	0.1252	C13 (c3)	0.1066
O5 (os)	-0.3689	O6 (os)	-0.3288
S1 (ss)	0.2090		