Molecular Dynamics Simulations, Challenges and Opportunities: A Biologist's Prospective

Article *in* Current Protein and Peptide Science · June 2017

DOI: 10.2174/1389203718666170622074741

4 authors, including:

Indu Kumari
National Institute Pathology Delhi
15 PUBLICATIONS **40** CITATIONS
SEE PROFILE

Mushtaq Ahmed
26 PUBLICATIONS **88** CITATIONS
SEE PROFILE

Yusuf Akhter
Babasaheb Bhimrao Ambedkar University
109 PUBLICATIONS **909** CITATIONS
SEE PROFILE

Molecular Dynamics simulations, challenges and opportunities: a biologist's prospective

Indu Kumari[1], Padmani Sandhu[2], Mushtaq Ahmed[1] and Yusuf Akhter[2*]

[1]Department of Environmental Science, School of Earth and Environmental Sciences, Central University of Himachal Pradesh, Shahpur, District-Kangra, Himachal Pradesh 176206, India.

[2]Centre for Computational Biology and Bioinformatics, School of Life Sciences, Central University of Himachal Pradesh, Shahpur, District-Kangra, Himachal Pradesh 176206, India.

***Correspondence:** Dr. Yusuf Akhter, Structural Bioinformatics Group, Centre for Computational Biology and Bioinformatics, School of Life Sciences, Central University of Himachal Pradesh, Shahpur, District, Kangra, 176206, Himachal Pradesh, India, Tel: +91-1892-237288 ext 309, 310, 311. Fax: +91-1892-2372286. Email: yusuf@daad-alumni.de; yusuf.akhter@gmail.com

Abstract

Molecular dynamics (MD) is a computational technique which is used to study biomolecules in virtual environment. Each of the constituent atoms represents a particle and hence the biomolecule embodies a multi-particle mechanical system analyzed within a simulation box during MD analysis. The potential energies of the atoms are explained by a mathematical expression consisting of different forces and space parameters. There are various software and force fields that have been developed for MD studies of the biomolecules. MD analysis has unravelled the various biological mechanisms (protein folding/unfolding, protein-small molecule interactions, protein-protein interactions, DNA/RNA-protein interactions, proteins embedded in membrane, lipid-lipid interactions, drug transport etc.) operating at the atomic and molecular levels. However, there are still some parameters including torsions in amino acids, carbohydrates (whose structure is extended and not well defined like that of proteins) and single stranded nucleic acids for which the force fields need further improvement, although there are several workers putting in constant efforts in these directions. The existing force fields are not efficient for studying the crowded environment inside the cells, since these interactions involve multiple

factors in real time. Therefore, the improved force fields may provide the opportunities for their wider applications on the complex biosystems in diverse cellular conditions. In conclusion, the intervention of MD in the basic sciences involving interdisciplinary approaches will be helpful for understanding many fundamental biological and physiological processes at the molecular levels that may be further applied in various fields including biotechnology, fisheries, sustainable agriculture and biomedical research.

Key words: Biomolecules; Force field; Ligand; Interactions; Molecular dynamics; Molecular mechanics

Abbreviations: AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Simulation), Correction energy map (CMAP), cyclic AMP (cAMP), Double Electron Electron Resonance (DEER), dummy spin-labels (DS), Glutamine-binding protein (GlnBP), Kinase-inducible domain (KID), Major histocompatibility complex (MHC), Molecular dynamics (MD), Molecular mechanics (MM), Post-translational modifications (PTM), Quantam mechanics (QM), Residue specific force field1 (RSFF1), Root mean square deviations (RMSD), T cell receptors (TCRs), Visual Molecular Dynamics (VMD).

Index of contents

1. Introduction
2. Different force fields for studying Biomolecules
3. Points to be taken care of before embarking on a molecular dynamics experiment
4. Molecular dynamics of protein-ligand systems
    - 4.1. Protein-small molecule complex MD analysis
    - 4.2. Protein-protein interaction specificity and binding
    - 4.3. Behaviour of nucleic acid-binding proteins
    - 4.4. Post-translational modifications of proteins
5. MD analysis pertaining to lipid bilayer biomembranes
    - 5.1. Physiological properties of plasma membrane analyzed by MD
    - 5.2. Conformational changes and functional mechanism of the integral membrane proteins
6. Molecular dynamics of carbohydrates
7. Molecular dynamics of DNA/RNA
8. Limitations of current force fields
    - 8.1. Poor description of secondary structure of biomolecules by current force fields
    - 8.2. Available force fields are deficient to study crowded environments of the cell
    - 8.3. Need to develop parameters for post-translational modifications
9. Conclusions and perspectives

1. Introduction

Modern molecular biology experiments involve *in vitro* studies to analyse the biomolecules in isolation in a test tube, the site-directed mutagenesis and transport or interactions of small molecules with the proteins. These investigations record the changes occurring at macro time scale but the microscopic details that lead to these macro level changes still remain a mystery. The Molecular dynamics (MD) of these biosystems in a simulation box provides a virtual microenvironment for analysing them in minute details. MD simulations with atomistic, physical parameters-based force fields involve principles of thermodynamics that may offer the potentials to gain new insights into the functional mechanisms operating at the level of single biomolecule to the biosystems (at the organeller levels) that are time and cost effective as compared to the conventional molecular gene expression, enzyme assays, protein isolation, characterization and crystal structure determination (Fig. 1) [1]. Several sampling methods are used for performing MD experiments on the biomolecules. These methods include adaptive umbrella sampling, integrated tempering sampling, metadynamics, accelerated MD and replica exchange, where MD has provided more robust conformational analysis possible for any of the biomolecules [2]. MD is based on the movement of biomolecules in the solvent box which follow Newtonian laws of motion. The spatial distribution of atomic properties of biomolecules is defined by force fields (FF) and its accuracy is essential for the success of an MD simulation experiment. All the atoms of a biomolecule (protein, lipid membrane, bound metal ions, associated water etc.) analyzed in an MD experiment are represented in the form of actual coordinates which generates new set of coordinates for the particles in terms of the energy of the system [3]. The parameters for force fields are developed by *ab-initio* calculations [4, 5]. The parameters of the FFs are adjusted in the light of experimental data obtained from neutron, X-ray and electron diffraction and scattering, NMR, infrared, Raman and neutron spectroscopy, etc. are used to generate the parameters for the atoms of the proteins. These parameters may also be obtained from *ab initio* or semi-empirical approaches of quantum mechanical calculations [6]. The atoms of biomolecules or a biosystem may be represented as simple elastic (harmonic) forces and FF replace these potentials of the system by a simplified model that is valid in the region being simulated [7,8]. However a typical expression for a FF may look like the following [9]:

$$E_{total} = \sum_{total} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\varphi - \gamma] + \sum_{i<i} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right].$$

Different mathematical symbols denote the various potential energy functions (E_{total}) which include equilibrium terms and force constants. Equilibrium term r-r_{eq} describes the energies of deformation of the bond length and θ-θ_{eq} represents the bond angle. The third term represents rotations around the chemical bond that are characterized by periodic energy terms (with n dihedral multiplicity, designated as n and steric hindrance defined by V_n), Φ is the dihedral angle and γ is the phase during rotation around the dihedral angle. The fourth term describes the van der Waals interatomic forces in the form of the Lennard–Jones (L-J) 12-6 potential, and the last term represents the Coulombic electrostatic potential. Some effects due to specific charge environments can be included using properly adjusted partial charges qi as well as the van der Waals parameters (Aij and Bij) and force constants i.e. K_r: bonds; K_θ: angle (Fig. 2) [10]. Here, the first four terms are referred to the local contributions (intramolecular) of the total energy within the molecules (bond stretching, angle bending and dihedral & improper torsions), and the last two terms describe the repulsive and Van der Waals interactions (in this case using a 12-6 L-J potential) and the Coulombic forces respectively (Fig. 2). All these parameters for FF are defined earlier and the parameters for the side chains of the biomolecules are determined at the end of the simulation experiment [11]. It is observed that the current FFs overestimate the α-helical fraction of the protein when compared to the experimental J-couplings [7]. Therefore, the accurate balance between the secondary structure propensities of protein is a central priority for the FF development for the protein because it may be considered as the readout for the protein folding, aggregation, association, and conformational changes which may be studied by MD. AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Simulation) FF for proteins has been developed to study permanent electrostatic multipole moments through the quadrupole at each atom which explicitly treats polarization effects in various chemical and physical environments [12]. However, there is need to further improve AMOEBA FF should also consider short-range electrostatic and Vander Waals components and their coupling to valence interactions that may improve this polarizable FF further. The improved FFs may use multicentre atomic charges and addition of the polarizability for backbone hydrogen bonds. Therefore, it would be necessary to change the balance of interactions occurring in the water model, which would be affected by any further changes. To keep it constant, it would be effective to introduce parallel terms for protein-solvent and solvent-solvent interactions [13]. Binding of two compounds/biomolecules follows the fundamental principles of physics i.e. molecular diffusion and intermolecular forces. However, the processes that occur inside the cells consist of many components of the cell in a single chemical reaction (isomerisation, ligation, hydrolysis, synthesis, transport etc.) [10,14]. A simple chemical reaction of one enzyme and substrate that occurs inside the cells becomes further complicated by the internal dynamics of the enzyme

(solvent accessible area, flexible and stable regions of the enzyme and the location of the binding site on the enzyme). The changes occurring in the dielectric properties of the proteins and at interfaces are poorly understood and the complex cellular environment possesses large changes due to the very large solute concentrations inside the cells [1,15]. It is important to note that the aim of this work is not to review all the available literature, in the traditional way. The motto of this article is rather to explore the interdisciplinary areas of computational biology and bioinformatics and their applications in other allied fields namely, biophysics and biochemistry from the view of a researcher of basic biology. This work may also provide a better understanding of the concepts of the MD simulations of the biomolecules protein, nucleic acid (DNA/RNA) [16] membrane lipids [17], carbohydrates, metal ions, nanoparticles etc. at the atomic levels and to further understand the possible mechanisms operating inside the cells [14,18-22] (Fig. 1). However, in contrast to the wealth of structural data (>1, 23,000 structures in the Protein Data Bank as retrieved on 10/16/2016), accurate information on different biomolecules related to their behaviour and dynamics is much more sparse. Therefore, the development and validation of simplified MD protocols and FFs still remains a challenging task.

2. Different force fields for studying Biomolecules

There are different FFs available to carry out MD studies. These include AMBER, CHARMM, GROMOS, and OPLS Force Fields [1,10,20,23]. There are also ongoing constant efforts to improve the existing FFs. These include the optimization of parameters for the biomolecules in terms of developing more robust mathematical expressions especially for the torsion angles between the two sets of three atoms having two atoms in common among them. Moreover, the integration of multiple time-steps based on dual Hamiltonian propagators will help to observe quantum mechanics/molecular mechanics (QM/MM) at a faster rate [24]. The AMBER package provides parameters for more composite molecules and fragments not currently available in the other FF libraries. Recently L-J parameters as well as the bond and angle parameters were refined for the General Amber Force Field. General Automated Atomic Model Parameterization was used to reparametrize the electrostatic and dihedral terms [25]. Yoo and Aksimentiev (2016) refined the L-J parameters describing non-bonded interactions between specific nitrogen-oxygen atomic pairs using CHARMM and AMBER FFs for amine-carboxylate, amine-phosphate, and aliphatic carbon-carbon interactions that provide remarkable equivalent results of MD simulations of proteins, nucleic acids and lipids with experiments [26]. The AMBER package released in 2012 also supports a wide range of FFs for the analysis of different molecules including the non-native variants of amino acids and the nucleic acids. Further it also included a

new fixed-charge with enhanced support for polarizable potentials. AMBER FF (ff99SB protein FF) was mainly meant for protein secondary structure balance and dynamics. The amino acid residues namely, isoleucine, leucine, aspartate and asparagine that exhibited particularly large deviations from the experimental distributions, indicated that the side chains of these residues were not modeled properly. Therefore, the improved version of AMBER ff99SB-ILDN has been proven to be better than that of the earlier ff99SB version [27,28]. However, new protein FF i.e. RSFF2 (residue-specific force field) based on protein coil library of 20 native amino acids was developed by modification of the AMBER ff99SB FF that has provided a better secondary structure balance for α-helical peptides, β-hairpins and Trp-cage, thereby, increasing the stability of α- helix and β-sheet simultaneously [2]. Recently Zgarbová et al. (2015) have refined the parameters for the sugar-phosphate backbone torsion β for AMBER FFs to improve the description of Z- and B-DNA (Fig. 2) [29]. The MD with explicit solvent was carried out including the refinement of the ε/ζ and glycosidic torsions. Further, sugar-phosphate backbone torsion angle β (β_{OL1}) was also refined so that it considered conformation-dependent solvation effects. This improved β_{OL1} was significantly able to model and describe the structure of the Z-DNA and its backbone as well as antiparallel guanine quadruplex (GDNA). It was slightly capable of describing the sub-states and also significantly described the sugar-phosphate backbone equilibria of B-DNA [29]. The accuracy of three current FFs in modeling the disordered peptides was assessed by combining enhanced-sampling MD simulations with NMR data. AMBER03w showed on an average more accuracy in describing the disordered systems. Moreover, it can be improved by incorporating the side-chains ILDN corrections that were developed for the AMBER99SB force field [28]. Best et al. (2012) have optimized the backbone correction energy map (CMAP) correction parameters with new side-chain parameters from complex quantam mechanics (QM) calculations for CHARMM36 [11]. The earlier available OPLS-AA/L FF reproduced potential mean forces well for some hydrophobic residues (such as Leu, Phe, Tyr, Trp, Val, Ile). However, it could not demonstrate the rotamer distributions (relative free energies of the three minima) well for Glu, His, Cys, Ser, Thr, and Asp. Therefore, the parameters shared between different χ torsions and different amino acid types (Fig. 1) i.e. χ_3 and parameters of non-hydrogen atoms were adjusted for better description of the rotational barriers in the improved version of the OPLS-AA/L FF i.e. residue specific force field1 (RSFF1). Though, RSFF1 provided a considerable balance between α-helical and β -sheet structures, it

also described well the transferability among different sequences [30]. The measurement of inter-structure distance provides the way for calculation of the flexibility and disordered regions of protein. This utility is present in GROMACS and provides a basis for multidimensional scaling and conformational clustering [31]. Moreover, there is further need of improvements in the current FFs, since there are several limitations in them which are discussed in the following sections.

3. Points to be taken care of before embarking on a molecular dynamics experiment

Before starting the MD studies on any of the biomolecules, we should be well acquainted with the literature available about the molecule of interest starting from the very beginning. Different factors should be considered such as availability of its tertiary structure in the PDB (Protein data bank), its existing monomeric and oligomeric states, biological unit and different conformational states should be taken into consideration. The conformational changes occurring in the binding site of the protein are stimulated by the action of additional inducers like co-factors [metal ions (Na^+, K^+, Mg^{2+} etc.), ATP, NADP etc]. If the protein is a crystal structure, then the structure which was refined at resolution better than 2.5 Å should be used [32] and all type of protein structures should be validated by Ramachandran plot based methods such as PROCHECK, MOLPROBITY etc. [33-35]. PROCHECK and MOLPROBITY use the distributions of χ side-chain torsion angles to validate the protein structure. The overall ‘‘clashscore,’’ which represents the number of bad all-atom clashes per thousand atoms (including hydrogen) should be less than five. The outlier amino acid residues should not be more than 5% in the Ramachandran plot [36]. The di-sulphide bonds in the proteins should be analyzed as these provide stability to the protein tertiary structures and play major role in the folding and unfolding of the proteins. The header of a PDB file contains information about the coordinates of atoms and quality of protein structure [resolution, completeness, I/σ (or signal-to-noise ratio) and the highest resolution shell] as well as indicators of the quality of the resulting structure, such as refinement parameters i.e. R-factor and R free obtained from refinement programs like Refmac or Phenix [37]. If the protein structure available in PDB is determined in solution using NMR spectroscopy, then, the data available in PDB about the *B*-factors of NMR models comprising of the root mean square deviations (RMSD) from average ensemble coordinate positions should be checked. The

biomolecular structures which show RMSD below 1.0 Å in the side-chain values, main-chain deviations should not be greater than 0.4 Å. The deviations in bond length and bond angle should be of the order of 0.018 Å and 1.1 Å respectively [38]. The chemical shifts, residual dipolar couplings and relaxation rates provided in NMR data indicate the quality of the protein structure [39,40]. Ideally, the non-protein atoms (HETATM, SOL and ions) should be removed from the PDB file. It should also be noted that protein structure is in its native form or a mutated version. If the structure represents a mutated version, then it can be modified to its native form using PyMOL/Coot [41.42]. The protein structure contains the whole biomolecule or if some parts of the biomolecule are missing, it should be checked and the missing parts should be repaired. Since the interactions of the biomolecules are described as bonded and non-bonded interactions and are defined in a set of parameters, therefore, it is necessary that all the atoms (α-carbon, polar donor, polar acceptor, apolar, and ionic) of the biomolecule should be explained by the empirical parameters of the FFs and there should not be any unusual behaviour of the atoms or steric clashes. It can be examined by visualizing the protein structure with graphical visualizers like Coot or PyMol that can check the amino acid residues which are lying outside the allowed regions of Ramachandran plot, and can bring them to the allowed regions by refining the φ and ψ angles of the outlying amino acid residues [41.42]. The overall charge of the system should be neutralized before the energy minimization. The water model in which protein has to be immersed should be selected cautiously as the loop and fold regions behave differently and the folding/unfolding processes affect both the equilibrium and dynamic properties of the system [43]. It can be concluded on the basis of the comparative structural behaviour of protein in different water models [44]. Among these, A99 and C22 of TIP4P/2005 were recently ranked as the most accurate force-fields for observing beyond microsecond time scale analysis [43]. For carrying out MD experiments, energy minimization, equilibration and production runs are usually carried out in steps as depicted in the flowchart (Fig. 3). GROMACS is one of the freely available suites which can be used and MD simulation is an effective tool which may provide crucial information about protein-ligand, protein-protein, post-translational modifications of proteins, protein-DNA/RNA and protein-lipid (embedded in membrane) interactions [45]. GROMACS provides different utilities like g_energy, g_rms, g_rmsf, g_hbond, g_dist, g_sas, g_density, g_isd etc to obtain this information. These are used to analyze the trajectory files obtained at the end of MD sessions and crucial conclusions could be drawn. The trajectory file

obtained at the end of MD should be analyzed by any molecular viewer such as Visual Molecular Dynamics (VMD) and if the native structure of the biomolecule distort/break during the simulations, it indicates the structure of biomolecule is not stable [46]. In such a situation, there is need to check every step and the parameters applied on the biomolecule and similar studies conducted on same biomolecules should be studied before carrying out the simulations again. All points described above should be kept in mind before proceeding for an MD experiment. Further, the duration of MD will depend on the type of biological system under the study. In case of molecular dynamics of protein, it depends on the length of amino acid sequence. Other factors, which may be taken into consideration for deciding the simulation time, may be folding-unfolding status and effect of mutation on the protein structure etc. For instance, protein-small molecule interactions of an average size of globular receptor can be analyzed well in 100 ns duration of MD session [47,48]. For another case, a membrane efflux protein was analyzed, the transport of ligand across the membrane through transporter protein is relatively longer phenomenon to capture, in such situation, the analysis which could be computationally less expensive and serve the purpose have been used by the researchers in smaller time scale to show only the direction of movement of the ligand towards the extracellular space after it was picked by the binding site at the intracellular side [49]. For MD studies of protein-protein interactions, the interacting sites at the interface of the protein-protein complex should be selected on the basis of already published and experimentally validated data. The potentially interacting proteins may be analyzed for the predicted disordered regions. It has been well established that, disordered regions, loops and turns are prone areas of molecular interactions between any two interacting partners [50]. There are online servers like Frustratometer, which compute the possible sites for protein-protein interactions as readout of frustrated structural parts of the proteins [51,52]. Some of the specific applications have been discussed in the following sections in detail.

4. Molecular dynamics of protein-ligand systems

4.1. Apo protein and protein-small molecule complex MD analysis

MD of protein in apo form and protein-ligand complex may be carried out at the atomic levels e.g. in the case of apo protein, the long MD analysis provided atomistic detail about different

folding states of the protein. Further, the enzyme catalyzed reactions where proteins interact with small molecules which cannot be observed by the conventional *in vitro* approaches like X-ray crystallography was demonstrated by MD analysis [55, 56]. The present FFs are well developed for studying the various functions of biomolecules (catalysis, ligand-binding locations, transport, motility, signal transduction, allosteric regulation, and molecular recognition) (Fig. 4). MD studies are widely used for the prediction of stability of the potential mutants of a protein molecule with higher catalytic efficiency or lower binding affinity for the toxins/molecules that may interfere with their normal function [35, 57, 58]. The modifications in the transporter protein's conformation by the normal motion patterns caused by the surrounding solvent or by the binding of a ligand molecule may be correlated with the efflux mechanism. These changes in the structural conformation of the transporter caused by the ligand binding are well studied by MD approach [14, 49]. A comparative MD simulation study on the NS3 protease of Hepatitis C virus has illustrated that the catalytic triad presents a dynamic behaviour which triggers a conformational instability in the protein that finally leads to abortion of the drug binding against the virus [59]. MD analysis of mutated HIV-1 IN protein (trans-activating regulatory protein**)** of the virus complexed with L-731,988 (β-diketo acids) could reveal the binding modes of this ligand and also helped in understanding the mechanism of its inhibition [60]. Natarajan and Senapati (2012) demonstrated the interactions and binding of the drugs (taxol and epothilone) with the mutants of tubulin dimer protein using MD simulations [61]. It was observed that the mutations primarily modified the M loop conformation of the protein primarily, enlarged the pocket volume and affected the distal sites of tubulin involved in the microtubule building processes [61]. The enzyme inhibitor process was reconstructed by MD simulations of free ligand binding that involves the placement of a ligand at certain distance from the protein of interest. The ligand binds first by diffusion, followed by specific interactions with protein. It binds to one or more sites in the protein. Benzamidine binding to trypsin showed various states during MD simulations [62]. Therefore, MD based studies have accelerated the research progress in the field of biomedical research by providing easy, fast and quantitative method for analysing the protein-drug interactions. The MD has helped in deciphering the behaviour of the drug in the binding pocket of the target protein, its stability and specificity. Hong et al (2016) have employed the integrated approach involving neutron spin echo (NSE) spectroscopy and MD to elucidate the collective movement of the domains which provide access to the catalytic site [64].

Therefore, the structural studies of protein and protein-small ligand molecule which will involve integrated strategies of *in vitro* and *in silico* may lead to better results [53, 54].

4.2. Protein-protein interaction specificity and binding

The interactions between proteins occur mainly by diffusion, non-covalent forces and the conformational flexibility present in the proteins. These are intertwined in such a way that they may be elucidated by MD simulations at atomistic levels. The non-covalent interactions occurring between biomolecules are the foundation of cellular biochemistry. The binding of two protein molecules is convoluted by complex internal dynamics with constantly changing dielectric properties inside and around the protein molecules. It is necessary to quantitatively access the binding region of the protein receptor and ligand so that the particular region of the protein involved in the interactions should be considered for MD as described in this article in earlier section. Moreover, the redistribution of solvation energy among the proteins and at the interfaces is poorly understood but they may be involved in protein-protein interactions or in crowded environment of the cells that represent varied solvent accessible surface area and dynamic solute concentrations inside the cells. Many of the causative forces act simultaneously in protein-protein interactions, making it even more complex phenomenon to analyze when the new incoming molecules appear in the analysis frame, rotational reorientation and reorganization, and steps of conformational selection and/or either induced fit or lock-key model hypotheses of the potential complexes may be operational. The protein-protein interactions of the biotin−streptavidin complex have been studied *in vitro*, however the atomistic details about the binding, optical transitions, harmonic vibrations and absorption spectra was obtained by molecular mechanical methods [65]. MD simulations can provide details upto atomic levels with high accuracy in nanoseconds to microseconds time scale in protein-protein interactions. However, recent studies have clearly demonstrated the importance of incorporating electrostatics with multiple charges, variable dielectrics and polarizability in atomistic force fields, especially for the better reproduction of hydrogen bonds, water defects/effects, aromatic chain stacking, pi-pi interactions, ion-aromatic ring interactions and other non-bonded interactions. It is commonly observed that in MD studies with the recent versions of the AMBER, CHARMM and OPLS FFs, the unfolded proteins tend to become much compact in simulations as compared to the experimental data [66]. While it is reported by MD analysis that the protein-protein interaction

specificity and binding advances by superposition of conformational selection and induced fit starting at least at surface-to-surface distances of 10 to 20 Å. Abriata and Peraro (2015) demonstrated that the conformational selection slowly occurs first and shifts to induced fit as the proteins move in close proximity[1]. Then, antiparallel orientation is adopted by dipolar moments and finally intermolecular contacts are established [1]. Antigen processing and antigen-antibody interactions are most important processes in the adaptive human immune response. When signalling processes occurring at the cellular levels were studied by MD approaches, they gave an insight into the diverse binding behaviour of the major histocompatibility complex (MHC) with different peptides. Different MHCs bind to same peptide which is aided by the conformational changes caused by the binding of different T cell receptors (TCRs). The MD studies can be used to explore the interactions of TCR-peptide-MHC ternary complex, the structural behavior of MHCs without ligand and simulations of the MHC containing trans-membrane regions. Further steered MD of TCR bound to pMHC and structural refinement of the homology modeled MHC proteins by MD may help to decipher their diverse binding modes. The conserved binding pattern of TCR/pMHC could be due to co-evolution of key amino acids involved in the interactions which can be studied in detail using MD [21]. A chimeric subunit model vaccine was designed using structure based rational approach. Thermodynamic stability of the vaccine was enhanced by disulphide engineering and its conformational accuracy and stability was assessed using MD simulations [21]. The integrated strategy comprised of equilibrium NMR analysis of thermal unfolding and long MD simulations deciphered the folding of W protein of bacteriophage lambda (gpW). It showed two-state-like fast folding dynamics and cooperative equilibrium unfolding behavior. It was observed that a complex pattern of structural changes occur at the atomic level. These changes result in detailed network of residue–residue couplings that lead to cooperative folding emergence [67]. However, the integrated approach which involved Terahertz spectroscopy and MD study of aqueous peptides showed the depth of hydration water around proteins and demonstrated the protein-protein interactions [63]. Therefore, MD approaches integrated with crystal structural studies may be used to decipher the possible mechanisms involved in the protein-protein interactions, specificity generation and the conformational changes occurring at the interface of the two proteins. It may lead to the stimulation, propagation and completion of the cellular processes.

4.3. Behaviour of nucleic acid-binding proteins

DNA-binding proteins take part in various cellular and molecular processes ranging from DNA polymerization to DNA repair and proof-reading, transcription and DNA super-coiling regulation. The extra-ordinary specificity and efficiency of DNA-binding proteins present a major theoretical riddle due to their enormous variations in the length and amino acid sequence interacting the DNA molecules. This fact is further supported by studies on protein–DNA complex structures. Those have explained about the molecular basis of specificity in individual or highly related complexes. However, there are no specific signature sequences or motifs have been found for a universal or generic recognition code which may adequately explain the observations about the specificity of all DNA binding proteins [68]. The sliding movements of DNA-binding protein and an RNA-binding protein (Barnase) along the B-DNA were studied earlier using MD simulations [69]. It was recorded for the protein-DNA interactions, there is lack of bonded interactions and only electrostatic and H-bonding interactions are prevalent. It was also observed that the protein molecule's movement was bidirectional in agreement with NMR data [69]. The contribution of dynamic behaviour of a single amino acid chain for the identification of one or more base pairs of telomere repeat binding factors (TRF1 and TRF2) that may help in resolving the conflicting experimental data was also explained by MD studies [22]. It was demonstrated by MD that two arginine side chains oscillated and induced direct interactions between DNA bases, DNA back bone and the transcription factor SKN-1. Disappearance of the protein–DNA salt bridges was observed with sub-nanosecond timescales of MD [69]. The conformational transitions in cyclic AMP (cAMP) receptor protein from *Mycobacterium tuberculosis* while binding to DNA was also analyzed by MD and normal mode analysis. It demonstrated that different orientations between the cAMP-binding and DNA-binding domains arise due to significant movement in the residues just before the C-helix and a larger movement at the end of C-helix. MD enabled us to propose a universal mechanism of allostery-mediated DNA recognition of the CRP [70]. MD analysis of DNA structure in free and bound to Cren7 (chromatin protein) protein at different temperatures was carried out. It was observed that dsDNA was stable in certain range of temperatures. It was also observed that some residues of protein have significantly formed H-bond and hydrophobic contacts with the DNA which contributed to the stability of the molecule. It was also reported that DNA got transitioned from B-like to A-like form with an increase in temperature [71]. Similar results were also

reported for the stabilization of DNA by Sso7d (hyperthermophilic proteins) in *Sulfolobus solfataricus* and further, it was found that only eight amino acid residues (Lys7, Tyr8, Lys9, Lys22, Trp24, Val26, Met29 and Arg43) were significant for the formation of the binary complex with the DNA [72] (Chen et al., 2012b). The intercalation of small drugs into DNA was also elucidated using MD and a mechanism of "detrapping" paradigm was explained in which minor grooves of DNA are involved in binding and also acting as inhibitors for the functional DNA-protein complexes. While intercalator molecule binds to DNA, there is a small energy barrier to overcome which favors unwinding of the DNA [73]. Metadynamics studies have shown that drug-DNA interactions are complex processes that elaborate the navigation of many basins (target) by the drug to non-specific contacts before reaching the significant pocket/groove exhibiting high binding affinities [74]. The alteration in the nucleic acid structure can be mainly of three types in DNA-protein complexes namely, DNA structure is not altered, the duplex integrity is consistent but DNA structure is disrupted and finally overall duplex organization is lost. The DNA Y-family polymerase IV (DPO4) protein of *Sulfolobus solfataricus* undergoes conformational change from an "open to close" state while interacting with DNA. MD analysis showed that DPO4 transitioned between multiple states while interacting with DNA and conformation distribution varied at different binding stages that regulated the DNA recognition [75]. MD was performed on the consensus sequence of DNA present in the (tumor suppressor) p53-specific binding site which elucidated that binding of p53 to DNA caused a local DNA conformational change and apparent DNA bending [76]. The DNA-protein interactions may occur due to the formation of hydrogen-bonds specifically between the protein and nucleotide bases, grooves mediated electrostatic forces (depending on the geometry and sequence of DNA/protein) and conformational changes due to sequence specificity [77-80]. Trichothecene molecules were reported to inhibit protein synthesis and MD analysis showed that trichodermin interacted with the nucleotide residues of 25S rRNA that constitute the petidyltransferase centre (PTC). However, the trichodermin resistance protein (60S ribosomal protein L3) was reported to overcome the inhibitory effects of trichothecene molecules. Further, MD of trichodermin resistance protein and 25S rRNA consisting of PTC showed that the W-finger region of the protein moved towards 25S rRNA and blocked the binding pocket of the trichodermin [81]. The large conformational change caused by binding of U1 (nuclear ribonucleoprotein) protein with RNA was investigated using MD. It was observed that there are modifications in hairpin and

internal loop of the protein upon binding to the RNA [82]. Therefore, MD elucidated the mechanisms and conformational changes induced by the DNA-binding proteins at the particular binding site or the minor/major groove of DNA/RNA involved in the cellular functions.

4.4. Post-translational modifications of proteins

The post-translational modifications (PTM) (addition or removal of chemical groups like phosphates) occur at the side chains of Ser, Thr, and Tyr in all organisms. There are several enzymatic and non-enzymatic PTMs. More than 250 distinct PTM reactions or 110 non-redundant PTM amino acids and protein termini have been reported [83]. The charge carried by phosphate is predominantly -2 at relevant pH of the cell. This results in large changes in electrostatic interactions that direct the energy landscapes influencing the folding of the protein to protein interactions and interactions with ligands that may also modulate its catalytic activities[84]. Groban et al. (2006) described the changes in the free energy landscape caused by phosphorylation of CDK2 which leads to conformational change in the activation loop of the protein using MD analysis [85]. The effects of phosphorylation could be observed on the immediate adjoining local structures of the protein. The phosphorylation like PTM may lead to repositioning of the interaction network in adjoining regions of the protein. There are some reports which showed that the conformational differences in the active and inactive forms of kinase were observed by MD simulations performed on Hck, an Src family kinases [86]. It is documented that few proteins contain hyperphosphorylated peptide regions i.e. various residues phosphorylated in close proximity e.g. hyperphosphorylation has been reported in the microtubule-associated protein tau which is commonly found in Alzheimer's disease [87]. The effect of hyperphosphorylation on the structural behavior of proteins is not well elucidated. Hamelberg et al. (2007) demonstrated that the hyperphosphorylation induced unfolding of the α-helix that leads to the formation of a compact 'arginine-claw' structure [88]. A single phosphate coordinates with five to six arginine residues while the other phosphates are solvent-exposed, a characteristic of arginine-claw. The folding mechanism of Kinase-inducible domain (KID) (transcriptional activator) was elucidated by the MD approach which showed that KID transforms a disordered protein to well folded protein by the changes induced by phosphorylation after post-translational modification [89]. MD analysis of ubiquitin protein showed long-range correlated motions which linked distal residues that resulted in the binding of

the interface of a protein. Further, this insinuated a mechanism which may lead to conformational selection, molecular recognition and biological function of ubiquitin [90]. It is demonstrated that myristoylation resulted from the interaction of myristate to N-terminal glycine residue [91]. It is shown that many proteins of eukaryotes are acetylated at the N-terminal residues namely, Ala, Ser, Val, Thr and Cys. Histone deacetylases complexed with its inhibitor showed the interactions active site residues with inhibitors using MD analysis which regulate the gene transcription [92]. The parameters were developed for PTMs which have been able to reproduce experimental hydration free energies (HFEs), a measure of hydrophobicity, which is one of the most important properties of the amino acids that drives protein folding, ligand binding or protein-lipid interactions. The parameters were generated for GROMOS FF (45a3 and 54a7) for over 250 diverse enzymatic and non-enzymatic PTMs because these parameters were fitted to reproduce experimental HFEs. This tool is available online for studying PTM with MD (http://vienna-ptm.univie.ac.at/) [83,93]. Although not enough experimental studies are available on following seven PTMs namely, carboxylysine, homocitrulline, S-carbamoyl-cysteine, citrulline, 2-oxo-histidine, S-nitrosocysteine and pyruvic acid [83], however, other post translational modifications such as methylation, sulfonation, myristoylation, acetylation, ubiquitination, glycosylation and methylation of amino acids are yet to be explored by MD analysis [94]. Post-translational modifications have important roles to play in controlling numerous eukaryotic cellular functions such as regulation of gene expression, chromosome structure and cell signalling. Only some of the post-translational modifications such as phopshorylation, acetylation, and ubiqutination have been studied in detail. While the mechanisms involved in the signal propagation by these modifications are unknown and hence, more attempts using combined techniques of *in vitro* MD are warranted to unravel the unknown mechanisms.

5. MD analysis pertaining to lipid bilayer biomembranes

5.1. Physiological properties of plasma membrane analyzed by MD

Lipids are very crucial components of the biological membranes and vary in their structure and function. Their biological roles range from signalling and energy storage to regulation of membrane protein functions. Although, experimental research is not yet fully efficient to unravel

the biological phenomenon of lipid functions to a great extent because of persisting difficulties in handling and isolation however, these limitations may be overcome with the application of molecular simulation techniques. The effect of presence of cholesterol (condensing effect) on lipid molecules in the form of change in the order parameter of lipid tails was recorded with the help of MD simulations [95]. MD simulations have also been able to explain lateral self-diffusion coefficient of lipid bilayer with changes in temperature and variation in the composition of the model membranes [96]. Nowadays, atomistic scale simulation, which is able to simulate bigger patches of lipid bilayer over a larger timescale, is an advancement in the field of lipid MD analysis [97].

These lipid membrane simulations showing lipid rafts and cholesterol interactions with the other lipid molecules have greatly changed the traditional outlook of lipid membrane. It has explored many more novel dimensions of lipid-lipid and lipid-protein interactions and their role in subsequent cellular signalling and trafficking across the cellular membrane. The major obstacles in the area of the lipid MD simulations are the accurate imitation of the asymmetric lipid bilayers with varied lipid composition. The coarse-grained models developed in the recent years have solved these problems to some extent [95,98]. Though understating the molecular driving forces for lipid domain formation still requires much more chemical details. However, many new developments in polarisable lipid FFs are in progress, which appear to be in harmony with the experimental data like area per lipid, electron density profile and deuterium order parameters of lipid membranes that will hopefully provide a more realistic view of the lipid molecules present in the biomembranes [99]. Futhermore, Lindorff-Larsen et al (2016) studied the self-assembly of the peptides with the phospholipid DMPC using combination of coarse-grained MD simulations, size-exclusion chromatography , circular dichroism spectroscopy, small-angle X-ray scattering, static light scattering and UV-Vis spectroscopy [56]. It revealed the molecular interactions occurring between the protein and lipids. Therefore it is needed that the biophysical studies should be carried out by the integration of experimental approach and computational techniques.

5.2. Conformational changes and functional mechanism of the integral membrane proteins

The experimental techniques like the high-resolution X-ray structure studies cannot address the dynamic nature of conformational changes which may occur during the movement of membrane proteins. Therefore, *in silico* approaches have the potential to utilize the data from both the

structural and biochemical research and it may reveal the functional mechanisms involved [100,101]. MD studies of protein embedded into the lipid membranes should be carried out by considering the complexity of the system. There are well established FFs consisting of parameters for lipid membranes. The AMBER [20], CHARMM [102], GROMACS [45] and NAMD [103] suites are suitable to carry out MD simulation of membrane embedded proteins. MD of lipid membranes may unravel the behavior and dynamics of the membranes at the atomic and molecular levels, with different solvents, drugs, glycoproteins and transporter proteins that is not easily accessible in different biochemical experiments. It also helps in understanding the change in properties of lipid membranes due to interactions with diverse types of molecules [17]. The first step in MD of protein embedded into the lipid membrane is the insertion of the protein into the lipid bilayer which mainly depends on the nature of the side chains of amino acid residues making the interface between the protein and the lipid bilayer. Further, the depth of insertion of the embedded protein in the lipid bilayer could be measured [104]. The solvent accessibility and other functional parameters are usually more sensitive to the distance from the polar surface of the membrane and therefore the distance between the phosphate groups in the lipid molecules and the centre of mass of the transmembrane helix in the inserted proteins are also quantified. The command g_density on GROMACS can be used to make a density profile for the head-groups and tail groups of phospholipids. It will be helpful to calculate the depth required to insert the protein in the lipid membrane. Recently, Javanainen (2014) developed a new protocol for embedding the proteins into the lipid bilayer by pushing them into the lipid bilayer membrane by the application of a high lateral pressure on the system from its sides [104]. It does not distort the initial geometry of the lipid bilayer [105]. Till now, several integral membrane proteins have been analyzed using MD simulations for understanding their functional role. The arginine of the D/E/NRY motif in Rhodopsin family G protein-coupled receptors (GPCRs) is conserved and has its role in salt bridge formation and also in the maintenance of endoderm-1 (Tre1) structure as revealed by MD simulations [106]. The transport of various ions and other molecules across the lipid bilayer and the mechanism operating in ion transport have been described in detail at the atomic levels through MD simulations. These studies revealed that different mechanisms are involved in the transport of the sodium and chloride ions. The membrane adsorbed sodium ions diffuse across the interface while chloride ions move through the middle of the pore with the bulk water and a crossover from ion-specific permeation was

observed when the pore radius was increased [107]. MD can be used to decipher the solubility of the drugs and the pathway involved in their transport across the membrane. Free binding energies and the interactions of the drugs with the solute and membranes can be calculated and further, the intra- and intermolecular states of the drugs in the surrounding environment and the membrane can be characterized by MD analysis [108]. MD studies have revealed that the adaptor protein exhibited interdomain dynamics which is essential for the formation of the tripartite complex composed of an outer membrane porin protein, inner membrane pump and a periplasmic adaptor protein which is involved in the efflux of compounds from Gram-negative bacteria and induce drug resistance in the bacteria [49,109]. MD simulations can address physical principles underlying key membrane processes, for instance, the puzzle connected with channel high selectivity and present a pertinent question that how can the flexible and highly dynamic systems like proteins be functionally highly selective [110]? It is now documented that uptake and release of various substances at the cellular levels across the biological membranes in the organisms is affected by secondary-active transporters. These transporters have been reported to possess a fold having 'inverted repeat' motif consisting of five transmembrane helices arranged in two sets orientated oppositely with respect to the lipid bilayer. Based on MD analysis of the sodium-benzylhydantoin transporter protein, Mhp1, from *Microbacterium liquefaciens* a mechanism has been proposed for switching from the outward- to the inward- facing state which may lead to the inward movement of sodium-benzylhydantoin, that is primarily achieved by a stiff movement of some transmembrane helices compared to the rest of the protein [111]. The activity of agonistic molecules (carboxylic acids and non-carboxylic acids) on the G-protein coupled receptor (GPCR), GPCR40 was observed using MD. It demonstrated ligand binding mechanisms of different agonists and fluctuations in the position of these molecules docked inside the transmembrane domain of the protein embedded in the lipid membrane [112]. The binding behaviour of antagonistic molecule [ML10302 (4-amino-5-chloro-2-methoxy-benzoic acid 2-piperidin-1-yl-ethyl ester)] with the serotonin, h5-HT_4 receptor (GPCR) was investigated using MD. It revealed that the receptor formed key interactions with the ligand and water molecules were transported from the extracellular milieu towards the putative hydrophobic pocket [113]. There are several transporters involved in various cellular processes where the mechanism of transport remains still unknown. The synergistic hybrid applications of Double Electron Electron Resonance (DEER) data for molecular structure with molecular simulation

methods have been reported to be an efficient strategy to explore the conformational states in the membrane protein structure and function [114]. The restrained-ensemble and dummy spin-labels (DS) simulation methods were employed to obtain the desired refined structures using global ensemble-based restrained energies. The force spin-spin distance distribution and histograms were calculated from multiple-copy MD simulation to match those obtained from DEER. The DS simulation method is based specially on parameterized dummy spin-labels that mimic the dynamics of all-atom multiple time-steps, spin-label and provides distance distribution between a pair of spin-labels that can be directly compared to the distance distributions obtained from DEER to access the reliability or correctness of an already existing 3D model structure of the system [114]. The crystallography studies did not provide details about the actual open and closed conformational states of glutamine-binding protein (GlnBP) of *E. coli.* However, the combined strategy which involved NMR residual dipolar coupling analysis, MD simulations, and single-molecule florescence resonance energy transfer methods showed four metastable states co-existing for apo-GlnBP in solution [115]. Such hybrid and multi-technology assimilated strategies together with classical molecular mechanics may provide important insights into the conformational dynamics of ion channels and transporter proteins. These unravelled mechanisms can be elucidated by MD and may be helpful in understanding the cell signalling, transport of drugs across the biological membranes and may also provide the basis for combating pathogenic resistant strains of the microbes.

6. Molecular dynamics of carbohydrates

The structural organization of carbohydrates is different from polypeptides, proteins and nucleic acids, since they do not have well defined spectral regions that determine the secondary-structural motifs. Therefore, they do not have well-organized tertiary structures in solution. However, the conventional experimental methods such as X-ray crystallography have a long history of their application in research on carbohydrates. Yet, these methods provide a single three-dimensional model that cannot explain the dynamic properties and behaviour of carbohydrates [116-118]. Terahertz spectroscopy study showed that the long range hydration dynamics of anti-freeze glycoproteins lead to the anti-freezing activity in these protein [119]. MD simulation has explored the multiple conformational states of oligosaccharides and glycoproteins. The FFs available for carbohydrates are GLYCAM06, CHARMM36 and

GROMOS 45A4. Various compounds can be studied by GLYCAM06, which is compatible with the AMBER MD package [120]. MD data were reported on carbohydrates and these include carbohydrate hydration, interaction of carbohydrates with ligands/other components, glycoproteins and other complexes of carbohydrates [121] (Fig. 4). Botulinum Neuro Toxin/B complexed with GD1A oligosaccharide and BoNT/B-GD1B complex was studied by MD which revealed different single binding mode of BoNT/B to GD1A i.e. single binding mode for GD1A and two binding modes for GD1B in the binding pocket of BoNT/B [121]. The dynamic behaviour of mannitol, sorbitol, α-cyclodextrin, glucose, ribose and deoxyribose in aqueous solution has also been explored using MD [115,122-126]. The MD study of glycan chains and the glycoproteins were separately carried out which showed the conformational flexibility of glycan chains. It also provided the interactions of protein-carbohydrate and changes occurring after the intercalation of ligand at the atomic levels [126]. There is deficiency of well-resolved crystallographic coordinates of the carbohydrates. Therefore, it is not easy to measure the changes induced by the attachment of glycan chain and its conformation, structure and dynamics of the protein. Moreover, the multi-technology driven studies Raman optical activity spectra and MD and QM/MM approach carried out on D-glucuronic acid and N-acetyl-D-glucosamine that are constituents of hyaluronan [127]. These studies revealed the contribution of hydration interactions towards the conformational behaviour of these two carbohydrates [127]. MD simulation study of β (1→4)-linked D-aldohexopyranose disaccharides indicated that solvent-exposed intramolecular H-bonding in aqueous carbohydrates represents a minor (possibly adverse) conformational driving as well as steering force. The intramolecular H-bonding may be the cause of the dominance of secondary structure motif typical of cello-oligosaccharides [128]. The disordered hydrogen bonding present in cellulose was studied using neutron crystallographic techniques and quantum mechanical calculations to elucidate the complex nature of hydrogen bonding in cellulose [129]. Further, studies involving such integrated approaches will provide understanding of the functioning of carbohydrates. All these processes can be studied in details at the atomic levels by integrated hybrid structural biology approaches using crystallography, SAXS, cryoelectron microscopy and MD of protein structure containing glycan chains.

7. Molecular dynamics of DNA/RNA

MD studies of nucleic acids are not very popular for various reasons. For instance, nucleic acids

697 are highly charged, non-globular and in many cases the crystal structures are not available (only
698 1650 of DNA and1190 of RNA PDB entries on 10/16/2016). The high charge densities on
699 nucleic acids can be accurately simulated by polarizable FFs that are still in developmental phase
700 [130,131]. Therefore, it is of importance to carefully develop parameters for the solvent and
701 electrostatics [132]. Initial MD studies on nucleic acids revealed structural differences in DNA-
702 DNA, RNA-RNA and DNA-RNA hybrid duplexes [133]. RNA duplex shows conformational
703 switching in uracil-cytosine base-pairing mediated by water molecules [134]. In one
704 conformation water acts both as H-bond donor and acceptor. The H-bond with uracil N3-H3
705 donor site did not show much change during simulations while the H-bond between an OH group
706 of water and cytosine N3 acceptor site of cytosine was found to rotate during the time of
707 simulation. While other conformations were observed to be similar to water-mediated UC base-
708 pair reported earlier in the crystal structure of the 23 S rRNA sarcin/ricin domain [134]. MD
709 studies have revealed that RNA and DNA duplexes exhibit differential structural flexibilities for
710 RNA. This varied structural behaviour of the RNA may be due to the local opening events
711 occurring in the major groove of RNA while such incidence was not observed in the case of
712 DNA. The solvent-base interactions, intra-strand hydrogen bonding among the RNA strands and
713 improved rigid body motions of the RNA at the atomic levels is the cause of locally open
714 conformations [76]. RNA hairpins that are an important class of RNA-RNA recognition motifs
715 were found to have loop to loop interactions. MD of these RNA motifs revealed structural
716 dynamics in the cation-binding pockets and major grooves that were reported to be involved in
717 the dimerization of retroviral RNA genomes such as HIV-1 and Moloney murine leukemia virus
718 [135]. Chen and García (2013) documented the *de novo* folding in case of three hyperstable RNA
719 tetraloops[136]. The formation of non-canonical loop-stabilizing interactions among the RNA
720 loops is the reason of this hyperstability. This was reproduced and found to be similar to
721 experimental data during replica exchange MD simulations. This accuracy was obtained by
722 reducing the L-J σ-parameters for base–base interactions, by adjusting the base–water
723 interactions to achieve correct decomposition enthalpy and entropy for dinucleotides stacking
724 and by the optimization of torsion angles of RNA [136]. MD analysis of DNA could be carried
725 out for 0.1 μs using the latest version of CHARMM [80]. The current MD simulations use
726 explicit solvent and addition of counter ions (typically Na^+ or K^+) to achieve neutrality in the
727 system. The umbrella sampling MD has revealed interesting mechanisms involved in bending

and base pair openings that may lead to local distortions involved in DNA repairing and epigenetic imprinting [132,137,138]. The folding/ unfolding of small DNA molecules can be studied within the μs time scale [80].

8. Limitations of current force fields

8.1. Poor description of secondary structure of biomolecules by current force fields

There are constant efforts to improve the present FFs. However, it was reported that CHARMM27 cannot fold all the WW domains of amyloid β protein due to over-stabilization of α-helical structures [139]. The parameters of backbone potentials were adjusted to reproduce the experimental J-couplings of Ala5 and the α-helical content of a poly-Ala-based peptide measured from NMR for AMBER (AMBERff99SB and ff03) and CHARMM22 [140,141]. For developing an accurate FF for protein, it is critical to obtain the local conformational free energy surface of each amino acid residue without secondary structure constraint. The other important feature is the side-chain χ^1 rotamer distributions of residues other than Gly, Ala, and Pro. The FFs for carbohydrates are still under development due to their extended and dynamic structures. MD simulations were carried out on immunoglobulin-binding protein G (NuG2b) and other small proteins which showed that these occur between the native state and a highly collapsed, H-bonded species. However, it is reported that these small proteins show cooperativity during the folding process started from a denatured state that may be expanded and largely unstructured in the beginning even at the lower concentrations of denaturants. This excessive intramolecular H-bonding recorded in the simulations suggested that there is a need to change the FF parameters to achieve the balance between all the deriving players, H-bonding, backbone solvation, and non-bonded interactions [142]. The FFs for nucleic acids are developed to some extent and are still in evolving phase. The parameters for nucleic acids need be improved for refinement of torsion (sugar pucker and glycosidic torsion parameter), van der Waals parameters and for the point charges. CHARMM27 misjudges the groove asymmetry and probably base pairing stability in DNA (clearly this is the case in RNA) or underestimates twisted trajectories of B-DNA [143]. To obtain the conformational equilibrium for RNA is particularly challenging: it is single stranded and difficult to replicate the backbone substrates of many important motifs such as the sarcin-ricin loop, some tetraloops, tetranucleotides, reversed kink-turns, and ribozymes among others

[29]. MD analysis of lipid membranes has provided novel information about the functioning of biological systems such as lung surfactants and tear-fluid layer but still there is scope of improvement [17]. Therefore, the parameters of FFs need to be modified in such a manner so that they are not biased towards particular amino acid residues/secondary structure.

8.2. Available force fields are deficient to study crowded environments of the cell

The cell consists of different micro/macromolecules distributed in dynamically changing environment with different thermodynamic and kinetic properties of the proteins, carbohydrates, nucleic acids and other molecules. This intracellular accumulation of variety of biomolecules and inorganic molecules also affects the structural stabilities, intermolecular/intramolecular binding affinities and enzymatic rates of proteins. However, various structural biology methods are expensive in terms of resources and time, such as NMR or different spectroscopies, usually require samples with relatively high protein concentrations and purity. While, this is not the case *in vivo*, proteins are neither present in that high concentration nor exist in isolation but may remain crowded by numerous other biomolecules and ligands. The accuracy of classical MD simulations is not good enough to capture protein behaviour at high concentration as the current FFs parameters cannot describe the crowded microenvironment [144]. There is also need to develop parameters for water model, which seems to lack the strength required to quickly dissociate bad complexes or even prevent their formation [1]. However, recently Yu et al. (2016) have studied the protein-protein interactions, surface diffusion of secondary metabolites in two dimensional space and protein-ligand interactions in the bacterial cytoplasm by MD [145]. While this study has pointed towards the transferability and compatibility deficiency of the CHARMM FF for close molecular interactions under conditions where FF parameters have to be developed extensively and artefacts, e.g. over stabilization of protein-protein contacts, diffusion of secondary metabolites and inclusion of the cell membranes in the MD study. Therefore, scientific community working in this area needs to put substantial efforts to develop FFs with more specialized applications that can be helpful in studying the protein aggregates in a single simulation box or can analyze the effect of high ion concentration on protein. This should also include polarizable FFs that may be able to capture the redistribution of electrons around each atom in response to the dynamic changes in the environment around the molecules [131].

8.3. Need to develop parameters for post-translational modifications

There are several known post-translational modifications, phosphorylation being most common among them. The energetics of phosphorylated amino acids has been described by MD analysis which showed that the phosphorylated amino acids are involved in the hydrogen bond formation [146]. Since phosphate is a charged entity, it represents a crucial interacting site for molecular mechanics energy functions. The electronic polarizability cannot be treated by commonly used fixed charge FFs and is probably found as the significant feature for various portions of a biomolecule that may interact with a phosphate group. Yet, polarizable FFs can be used to analyze such interactions of biomolecules. However, post-translational modifications can modulate protein charge in the whole protein/specific manner and thus may also act as 'electrostatic switches'. There are different types of post-translational modifications such as sulfonation, lysine acetylation, and glutamate methylation, the last two being important for histone modifications [84]. Therefore, there is a need to develop new and efficient parameters for van der Waals interactions which will occur between the atoms of protein and the covalently attached groups during PTM. PTMs are involved in bond stretching, angle bending, dihedral & improper torsions between the atoms of the protein and the attached group of atoms which may be consisted of carbon, phosphorous, sulphur and other atoms. However, the potential energy functions for the bond angle, dihedral torsion and electrostatic interactions are well developed for proteins. Nevertheless, proteins interact with various newly discovered chemical groups implicated by PTMs like acetyl, pyruvic acid, carboxylysine etc., then, there is need to develop potential energy parameters for these changes which are not being explored using classical MD analysis and by experimental studies extensively yet.

9. Conclusions and perspectives

Recently, it has been reported that the current versions of popular FFs (including AMBER99SB-ILDN, CHARM22-CMAP, GROMOS 45a3 and 54a7, and OPLS-AAL) are not accurate enough for studying the proteins existing in crowded environments because of the excessive tendency of interactions [144]. The events involved in the folding of the proteins in their native forms can be elucidated by MD simulations with molecular mechanics based FFs [147]. However, the current FFs have problems in describing the discrepancies in the local conformational preferences (secondary structure) of different amino acid residues [141]. Therefore, there is a serious need to

improve the existing FFs and also to develop additional parameters for different forms of proteins, nucleic acids and carbohydrates. There is also need for the development of polarisable FFs for studying the biomolecules occurring in crowded environments. However, there is an urgent need for developing easy protocols/methods so that the simulations can be carried out without much difficulty. Although, in case of lipid membranes, the FFs are well defined for various types of lipids but still there is a need to significantly improve the parameters for the protein-lipid (polar-head/tail) interactions, solvent-lipid interactions and solute-lipid interactions. There are many known MD reports that involve transition metals as co-factors of proteins that play important roles in the biological processes. Although, there are parameters developed for some of the monovalent and divalent metal ions in different FFs, however, there are still some known metals for which new parameters need to be developed. There is also a striving need of automation in the MD protocols so that, the researchers not having technical background in high end computing may also be able to use these protocols for carrying out the analysis for studying important biological phenomena. Easy going methods for MD usage protocols may change the prevailing scenario preferentially highly technical users to a popular tool for use by the common biologist folk. If all this is achieved in the near future, MD analysis definitively will be able to venture its way into the main stream of research in the basic and applied biology. There will be an integration of highly specialized "dry" labs with strong theoretical background and evident quantitative skills with the general biologists. Moreover, in this way, the tools and technologies developed in the "dry" labs will provide new opportunities resulting into leading methodologies which will be utilised in the "wet" labs, often exist without such backgrounds, but working in crucial aspects of the basic biology on different biomolecules.

Acknowledgements

We acknowledge Central University of Himachal Pradesh and Bioinformatics Resources & Applications Facility, Centre for Development in Advanced Computing, Pune for providing the computational infrastructure. We acknowledge Ms. Shweta Thakur for helping us during the editing and final formatting of the manuscript.

Funding

IK receives research stipend from University Grant Commission, Govt. of India (UGC). PS

receives senior research fellowship from Indian Council of Medical Research, Govt. of India. Research in MA lab is funded by Science and Engineering Research Board, DST, Govt. of India (SERB) and UGC. Research in YA lab is supported by extramural research funds from UGC, ICMR and SERB.

References

[1] Abriata, L.A.; Peraro M.D. Assessing the potential of atomistic molecular dynamics simulations to probe reversible protein-protein recognition and binding. *Scientific Rep*., 2015, 5.

[2] Zhou, C.Y.; Jiang, F.; Wu, Y.D. Residue-specific force field based on protein coil library. RSFF2: Modification of AMBER FF99SB. *The J. Phys. Chem. B*, 2014, 119,1035-47.

[3] González, M.A. Force fields and molecular dynamics simulations. *École thématique de la Société Française de la Neutronique*. 2011, 12, 169-200.

[4] Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules *J. Am. Chem. Soc*., 1995, 117, 5179− 5197.

[5] Jorgensen, W.L.; Maxwell, D.S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. The Am. Chem. Soc*., 1996, 118(45), 11225-11236.

[6] Ball, K.A.; Wemmer, D.E.; Head-Gordon, T. Comparison of structure determination methods for intrinsically disordered amyloid-β peptides. *The J. Phys. Chem. B*. 2014, 118, 6405-6416.

[7] Best, R.B.; Buchete, N.V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys. J.*, 2008, 95, L07-9.

[8] Groenhof, G. Introduction to QM/MM simulations. In: Biomolecular Simulations: Methods and Protocols; Monticelli, L. Salonen E. Ed.; Springer, 2013, pp. 43-66.

[9] Levitt, M.; Lifson, S. Refinement of protein conformations using a macromolecular energy minimization procedure. *J. Mol. Biol*., 1969, 46, 269-79.

[10] Wang, W.; Donini, O.; Reyes, C.M.; Kollman, P.A. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-

protein, and protein-nucleic acid noncovalent interactions. *Ann. Rev. Biophys. Biomol. Stru.*, 2001, 30, 211-43.

[11] Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.M.; Mittal, J.; Feig, M.; MacKerellJr, A.D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. *J. Chem. Theory. Comp*., 2012, 8, 3257-3273.

[12] Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J.W.; Ren, P. Polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comp*., 2013, 9(9), 4046-4063.

[13] Best, R.B.; Hummer, G. Optimized molecular dynamics force fields applied to the helix− coil transition of polypeptides. *The J. Phys. Chem. B*., 2009, 113, 9004-15.

[14] Ferreira, R.J.; Ferreira, M.JU.; dos Santos, D. J.Insights on P-glycoprotein's efflux mechanism obtained by molecular dynamics simulations. *J. Chem. Theory Comp*., 2012, 8, 1853-64.

[15] Schaeffer, R.D.; Fersht, A.; Daggett, V.; Combining experiment and simulation in protein folding: closing the gap for small model systems. *Curr. Opin. Struc. Biol*.; 2008, 18, 4-9.

[16] Henke, P.S.; Mak, C.H. Free energy of RNA-counterion interactions in a tight-binding model computed by a discrete space mapping. *The J. Chem. Phys*., 2014, 141, 064116.

[17] Hirtz, M.; Kumar, N.; Chi, L. Simulation Modeling of Supported Lipid Membranes–A Review. *Curr. Top. Med. Chem*., 2014, 14, 617-23.

[18] Khandelwal, A.; Lukacova, V.; Comez, D.; Kroll, D.M.; Raha, S.; Balaz S. A combination of docking, QM/MM methods, and MD simulation for binding affinity estimation of metalloprotein ligands. *J. Med. Chem*., 2005, 48, 5437-47.

[19] Chatterjee, S.; Gersten, B.; Thakur, S.; Burin, A. Molecular dynamics simulations of a single stranded (ss) DNA. *Mol. Simu*., 2007, 33, 573-6.

[20] Salomon-Ferrer, R.; Case, D.A.; Walker, R.C. An overview of the Amber biomolecular simulation package. *Wiley Interdisci. Rev.: Comp. Mol. Sci*., 2013, 3(2), 198-210.

[21] Knapp, B.; Demharter, S.; Esmaielbeiki, R.; Deane, C.M. Current status and future challenges in T-cell receptor/peptide/MHC molecular dynamics simulations. *Brief Bioinform*., 2015, 16, 1035-44.

[22] Etheve, L.; Martin, J.; Lavery, R. Dynamics and recognition within a protein–DNA complex: a molecular dynamics study of the SKN-1/DNA interaction. *Nucl. Acids Res.*, 2016, 44, 1440-8.

[23] Mu, Y.; Kosov, D.S.; Stock, G. Conformational dynamics of trialanine in water. 2. Comparison of AMBER, CHARMM, GROMOS, and OPLS force fields to NMR and infrared experiments. *The J. Phys. Chem. B.*, 2003, 107(21), 5064-5073.

[24] Chen, Y.; Kale, S.; Weare, J.; Dinner, A.R.; Roux, B. Multiple Time-Step Dual-Hamiltonian Hybrid Molecular Dynamics–Monte Carlo Canonical Propagation Algorithm. *J. Chem. Theory Comp.*, 2016, 12, 1449-58.

[25] Boulanger, E.; Huang, L.; MacKerell Jr, A.D.; Roux, B. Improved Lennard-Jones Parameters for Accurate Molecular Dynamics Simulations. *Biophys. J.*, 2016, 110, 646a.

[26] Yoo, J.; Aksimentiev, A. Improved Parameterization of Amine-Carboxyate, Amine-Phosphate, and Aliphatic Carbon-Carbon Interactions for Molecular Dynamics Simulations using the Charmm and Amber Force Fields. *Biophys. J.*, 2016, 110, 646a.

[27] Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Pro.: Str. Func. Bioinfo.*, 2010, 78(8),1950-1958.

[28] Palazzesi, F.; Prakash, M.K.; Bonomi, M.; Barducci, A. Accuracy of current all-atom force-fields in modeling protein disordered states. *J. Chem. Theory Comp.*, 2014, 11, 7.

[29] Zgarbová, M.; Šponer, J.; Otyepka, M.; Cheatham, T.E. III.; Murillo, R.G.; Jurečka, P. Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z-and B-DNA. *J. Chem. Theory Comp.*, 2015, 11, 5723-36.

[30] Jiang, F.; Zhou, C.Y.; Wu, Y.D. Residue-specific force field based on the protein coil library. RSFF1: modification of OPLS-AA/L. *The J. Phys. Chem. B.*, 2014, 118, 6983-98.

[31] Connolly, T.G.; Ando, D.; Newsam, S.D.; Gopinathan, A.; Colvin, M.E. Universal Metrics of Interstructure Distance for Flexible and Intrinsically Disordered Proteins. *Biophys. J.*; 2016, 110, 644a.

[32] Rydzewski, J.; Jakubowski, R.; Nowak, W. Communication: Entropic measure to prevent energy over-minimization in molecular dynamics simulations. *The J. Chem. Phys.*, 2015, 143, 171103.

[33] Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. App. Crystall*., 1993, 26, 283-91.

[34] Davis, I.W.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucl. Acids Res*., 2004, 32, W615-9.

[35] Kumari, I.; Chaudhary, N.; Sandhu, P.; Ahmed, M.; Akhter, Y. Structural and mechanistic analysis of engineered trichodiene synthase enzymes from *Trichoderma harzianum*: towards higher catalytic activities empowering sustainable agriculture. *J. Biomol. Struc. Dy*., 2016a, 34, 1176-89.

[36] Read, R.J.; Adams, P.D.; Arendall, W.B. III.; Brunger, A.T.; Emsley, P.; Joosten, R.P.; Kleywegt, G.J.; Krissinel, E.B.; Lütteke, T.; Otwinowski, Z.; Perrakis, A.; Richardson, J.S.; Sheffler, W.H.; Smith, J.L.; Tickle, I.J.; Vriend, G.; Zwart, P.H. A new generation of crystallographic validation tools for the protein data bank. *Struc.*, 2011, 19, 1395-1412.

[37] Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *Febs. J.*, 2008, 275, 1-21.

[38] Rosato, A.; Tejero, R.; Montelione, G.T. Quality assessment of protein NMR structures. *Curr. Opin. Struc. Bio*., 2013, 23, 715–724.

[39] Montelione, G.T.; Nilges, M.; Bax, A.; Güntert, P.; Herrmann, T.; Richardson, J.S.; Schwieters, C.; Vranken, W.F.; Vuister, G.W.; Wishart, D.S.; Berman, H.M.; Kleywegt, G.J.; Markley, J.L. Recommendations of the wwPDB NMR validation task force. *Struc.*, 2013, 21, 1563-1570.

[40] Gutmanas, A.; Adams, P.D.; Bardiaux, B.; Berman, H.M.; Case, D.A.; Fogh, R.H.; Güntert, P.; Hendrickx, P.M.S.; Herrmann, T.; Kleywegt, G.J.; Kobayashi, N.; Lange, O.F.; Markley, J.L.; Montelione, G.T.; Nilges, M.; Ragan, T.J.; Schwieters, C.D.; Tejero, R.; Ulrich, E.L.; Velankar, S.; Vranken, W.F.; Wedell, J.R.; Westbrook, J.; Wishart, D.S.; Vuister, G.W. NMR exchange format: a unified and open standard for representation of NMR restraint data. *Nat. struc. Mol. Bio*., 2015, 22, 433-434.

[41] DeLano, W.L. The PyMOL Molecular Graphics System. DeLano Scientific LLC, San Carlos, CA. 2002.

975 [42] Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. Acta
976 Crystall. *Sec D: Biolog. Crystall*., 2004, 60, 2126-32.
977 [43] Palazzesi, F.; Salvalaglio, M.; Barducci, A.; Parrinello, M. Communication: Role of
978 explicit water models in the helix folding/unfolding processes. *The J. Chem. Phys*., 2016,
979 145, 121101.
980 [44] Gupta, M.; Nayar, D.; Chakravarty, C.; Bandyopadhyay, S. Comparison of hydration
981 behavior and conformational preferences of the Trp-cage mini-protein in different rigid-
982 body water models. *Phys. Chem. Chem. Phy*., 2016, 18(48), 32796-32813.
983 [45] Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.;
984 Smith, J.C.; Kasson, P.M.; Spoel, D.V.; Hess, B.; Lindah, E. GROMACS 4.5: a high-
985 throughput and highly parallel open source molecular simulation toolkit. *Bioinfo*., 2013,
986 29, 845-54.
987 [46] Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol.*
988 *Graph*., 1996, 14, 33-8.
989 [47] Agulló, L.; Malhotra, S.; Fissolo, N.; Montalbanb, X.; Comabella M. Molecular
990 dynamics and intracellular signaling of the TNF-R1 with the R92Q mutation. *J.*
991 *Neuroimmuno*., 2015, 289, 12–20.
992 [48] Wang, D.F.; Helquist, P.; Wiech, N.L.; Wiest, O. Toward selective histone deacetylase
993 inhibitor design: homology modeling, docking studies, and molecular dynamics
994 simulations of human class I histone deacetylases. *J. Med. Chem*., 2015a, 48, 6936-47.
995 [49] Sandhu, P.; Akhter, Y. The drug binding sites and transport mechanism of the RND
996 pumps from *Mycobacterium tuberculosis*: Insights from molecular dynamics simulations.
997 *Arch. Biochem. Biophys*., 2016, 592, 38-49.
998 [50] Papaleo, E.; Saladino, G.; Lambrughi, M.; Lindorff-Larsen, K.; Gervasio, F. L.;
999 Nussinov, R. The role of protein loops and linkers in conformational dynamics and
1000 allostery. *Chem. Rev*., 2016, 116(11), 6391-6423.
1001 [51] Rana, A.; Akhter, Y. A multi-subunit based, thermodynamically stable model vaccine
1002 using combined immunoinformatics and protein structure based approach. *Immunobio*.,
1003 2016, 221, 544-57.

[52] Jenik, M.; Parra, R. G.; Radusky, L. G.; Turjanski, A.; Wolynes, P. G.; Ferreiro, D. U. Protein frustratometer: a tool to localize energetic frustration in protein molecules. *Nuc. Acids Res.*, 2012, gks447.

[53] Harish, B.; Swapna, G.V.T.; Kornhaber, G.J.; Montelione, G.T.; Carey, J. Multiple helical conformations of the helix-turn-helix region revealed by NOE-restrained MD simulations of tryptophan aporepressor, TrpR. *Pro.: Str. Func. Bioinfo.*, 2017, 85(4), 731-740.

[54] Perilla, J.R.; Zhao, G.; Lu, M.; Ning, J.; Hou, G.; Byeon, I.J.L.; Gronenborn, A.M.; Polenova, T.; Zhang, P. CryoEM Structure Refinement by Integrating NMR Chemical Shifts with Molecular Dynamics Simulations. *The J. Phys. Chem. B.*, 2017.

[55] Lange, O.F.; Spoel, D.V.; de. Groot, B.L. Scrutinizing molecular mechanics force fields on the submicrosecond timescale with NMR data. *Biophys. J.*, 2010, 99, 647-55.

[56] Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Shaw, D.E. Picosecond to millisecond structural dynamics in human ubiquitin. *J. Phys. Chem. B.*, 2016, 120(33), 8313-8320.

[57] Pathania, S.; Randhawa, V.; Bagler, G. Prospecting for novel plant-derived molecules of *Rauvolfia serpentina* as inhibitors of Aldose Reductase, a potent drug target for diabetes and its complications. *PloS one*. 2013, 8, e61327.

[58] Wang, F.; Wan, H.; Hu, J.P.; Chang, S. Molecular dynamics simulations of wild type and mutants of botulinum neurotoxin A complexed with synaptic vesicle protein 2C. *Mol. BioSys.*, 2015b, 11, 223-31.

[59] Kramer, M.; Halleran, D.; Rahman, M.; Iqbal, M.; Anwar, M.I.; Sabet, S.; Ackad, E.; Yousef, M. Comparative molecular dynamics simulation of hepatitis C virus NS3/4A protease (genotypes 1b, 3a and 4a) predicts conformational instability of the catalytic triad in drug resistant strains. *PloS one*. 2014, 9, e104425.

[60] Brigo, A.; Lee, K.W.; Fogolari, F.; Mustata, G.I. J.M. Briggs, Comparative molecular dynamics simulations of HIV-1 integrase and the T66I/M154I mutant: Binding modes and drug resistance to a diketo acid inhibitor. *Pro. Struc. Func. Bioinfo.*, 2005, 59, 723-41.

[61] Natarajan, K.; Senapati, S. Understanding the basis of drug resistance of the mutants of αβ-tubulin dimer via molecular dynamics simulations. *PloS one*. 2012, 7, e42351.

[62] Buch, I.; Giorgino, T.; De. Fabritiis, G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *PNAS*. 2011, 108, 25.

[63] Sushko, O.; Dubrovka, R.; Donnan, R.S. Sub-terahertz spectroscopy reveals that proteins influence the properties of water at greater distances than previously detected. *The J. Chem. Phys.*, 2015, 142(5), 02B602_1.

[64] Hong, L.; Jain, N.; Cheng, X.; Bernal, A.; Tyagi, M.; Smith, J.C. Determination of functional collective motions in a protein at atomic resolution using coherent neutron scattering. *Sci. Adv.*, 2016, 2(10), e1600886.

[65] Bykhovski, A.; Zhang, W.; Jensen, J.; Woolard, D. Analysis of Electronic Structure, Binding, and Vibrations in Biotin–Streptavidin Complexes Based on Density Functional Theory and Molecular Mechanics. *The J. Phys. Chem. B.*, 2012, 117(1), 25-37.

[66] Piana, S., Larsen, K.L., Shaw, D.E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.*, 2011, 100, L47−L49.

[67] Sborgi, L.; Verma, A.; Piana, S.; Larsen, K.L.; Cerminara, M.; Santiveri, C.M.; Shaw, D.E.; Alba, de E.; Muñoz, V. Interaction networks in protein folding via atomic-resolution experiments and long-time-scale molecular dynamics simulations. *J. Amer. Chem. Soc*., 2015, 137, 6506-16.

[68] Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nuc. Acids Res.*, 2001, *29*(13), 2860-2874.

[69] Givaty, O.; Levy Y. Protein sliding along DNA: dynamics and structural characterization. *J. Mol. Bio*., 2009, 385, 1087-97.

[70] Kumar, P.; Joshi, D.C.; Akif, M.; Akhter, Y.; Hasnain, S.E.; Mande, S.C. Mapping conformational transitions in cyclic AMP receptor protein: crystal structure and normal-mode analysis of *Mycobacterium tuberculosis* apo-cAMP receptor protein. *Biophys. J.*, 2010, 98, 305-14.

[71] Chen, L.; Zhang, J.L.; Yu, L.Y.; Zheng, Q.C.; Chu, W.T.; Xue, Q.; Zhang, H.X.; Sun, C.C. Influence of hyperthermophilic protein Cren7 on the stability and conformation of

DNA: Insights from molecular dynamics simulation and free energy analysis. *The J. Phys. Chem. B.*, 2012, 116, 12415-25.

[72] Chen, L.; Zheng, Q.C.; Yu, L.Y.; Chu, W.T.; Zhang, J.L.; Xue, Q.; Zhang, H.X.; Sun, C.C. Insights into the thermal stabilization and conformational transitions of DNA by hyperthermophile protein Sso7d: molecular dynamics simulations and MM-PBSA analysis. *J. Biomol. Struc. Dyn*., 2012b, 30, 716-27.

[73] Mukherjee, A.; Lavery, R.; Bagchi, B.; Hynes, J.T. On the molecular mechanism of drug intercalation into DNA: a simulation study of the intercalation pathway, free energy, and DNA structural changes. *J. Amer. Chem. Soc*., 2008, 130, 9747-55.

[74] Vargiu, A.V.; Ruggerone, P.; Magistrato, A.; Carloni, P. Dissociation of minor groove binders from DNA: insights from metadynamics simulations. *Nucl. Acids Res*., 2008, 36, 5910-21.

[75] Chu, X.; Liu, F.; Maxwell, B.A.; Wang, Y.; Suo, Z.; Wang, H.; Han, W.; Wang, J. Dynamic conformational change regulates the protein-DNA recognition: an investigation on binding of a Y-family polymerase to its target DNA. *PLOS Comput. Biol*., 2014, 10, e1003804.

[76] Pan, Y.; MacKerell Jr, A.D. Altered structural fluctuations in duplex RNA versus DNA: a conformational switch involving base pair opening. *Nucl. Acids Res*., 2013, 31, 7131-40.

[77] Joshi, R.; Passner, J.M.; Rohs, R.; Jain, R.; Sosinsky, A.; Crickmore, M.A.; Jacob, V.; Aggarwal, A.K.; Honig, B.; Mann, R.S. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell*. 2007, 131, 530-43.

[78] Tolstorukov, M.Y.; Colasanti, A.V.; McCandlish, D.; Olson, W.K.; Zhurkin, V.B. A novel roll-and-slide mechanism of DNA folding in chromatin: implications for nucleosome positioning. *J. Mol. Biol*., 2007, 371, 725-38.

[79] Balasubramanian, S.; Xu, F.; Olson, W.K. DNA sequence-directed organization of chromatin: structure-based computational analysis of nucleosome-binding sequences. *Biophys. J*., 2009, 96, 2245-60.

[80] Pérez, A.; Luque, F. J.; Orozco, M. Frontiers in molecular dynamics simulations of DNA. *Acc. Chem. Res*., 2011, 45, 196-205.

[81] Kumari, I.; Ahmed, M.; Akhter, Y. Deciphering the protein translation inhibition and coping mechanism of trichothecene toxin in resistant fungi. *The Intl. J. Biochem. Cell Bio*., 2016b, 78, 370-376.

[82] Reyes, C.M.; Kollman, P.A. Structure and thermodynamics of RNA-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. *J. Mol. Biol*., 2000, 297, 1145-58.

[83] Petrov, D.; Margreitter, C.; Grandits, M.; Oostenbrink, C.; Zagrovic, B. A systematic framework for molecular dynamics simulations of protein post-translational modifications. *PLoS Comput. Biol*., 2013, 9, e1003154.

[84] Narayanan, A.; Jacobson, M.P. Computational studies of protein regulation by post-translational phosphorylation. *Curr. Opin. Struc. Biol*., 2009, 19, 156-63.

[85] Groban, E.S.; Narayanan, A.; Jacobson, M.P. Conformational changes in protein loops and helices induced by post-translational phosphorylation. *PLoS Comp. Bio.*, 2006, 2, e32.

[86] Banavali, N.K.; Roux, B. Anatomy of a structural pathway for activation of the catalytic domain of Src kinase Hck. *Pro. Str. Func. Bioinfo*., 2007, 67, 1096-1112.

[87] Mazanetz, M.P.; Fischer, P.M. Untangling tau hyperphosphorylation in drug design for neurodegenerative diseases. *Nat. Rev. Drug Dis*., 2007, 6, 464-79.

[88] Hamelberg, D.; Shen, T.; McCammon, J.A. A proposed signaling motif for nuclear import in mRNA processing via the formation of arginine claw. *Proc. Nat. Acad. Sci*., 2007, 104, 14947-51.

[89] Chen, H.F. Molecular dynamics simulation of phosphorylated KID post-translational modification. *PLoS One*. 2009, 4, e6516.

[90] Fenwick, R.B.; Martín, S.E.; Richter, B.; Lee, D.; Walter, K.F.A.; Milovanovic, D.; Becker, S.; Lakomek, N.A.; Griesinger, C.; Salvatella, X. Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition. *J. Amer Chem. Soc*., 2011, 133, 10336-9.

[91] Cordeddu, V.; Schiavi, E.D.; Pennacchio, L.A.; Ma'ayan, A.; Sarkozy, A.; Fodale, V.; Serena, C.; Cardinale, A.; Martin, J.; Schackwitz, W.; Lipzen, A.; Zampino, G.; Mazzanti, L.; Digilio, M.C.; Martinelli, S.; Flex, E.; Lepri, F.; Bartholdi, D.; Kutsche, K.; Ferrero, G.B.; Anichini, C.; Selicorni, A.; Rossi, C.; Tenconi, R.; Zenker, M.; Merlo, D.;

Dallapiccola, B.; Iyengar, R.; Bazzicalupo, P.; Gelb, B.D.; Marco, T. Mutation of SHOC2 promotes aberrant protein N-myristoylation and causes Noonan-like syndrome with loose anagen hair. *Nat. Gen*., 2009, 41, 1022-6.

[92] Wang, D.F.; Helquist, P.; Wiech, N.L.; Wiest, O. Towards selective histone deacetylase inhibitor design: homology modeling, docking studies, and molecular dynamics simulations of human class I histone deacetylases. *J. Med Chem*., 2005, 48(22), 6936-6947.

[93] Margreitter, C.; Petrov, D.; Zagrovic, B. Vienna-PTM web server: a toolkit for MD simulations of protein post-translational modifications. *Nucl. Acids Res*., 2013, 41, W422-6.

[94] Sprung, R.; Chen, Y.; Zhang, K.; Cheng, D.; Zhang, T.; Peng, J.; Zhao, Y. Identification and validation of eukaryotic aspartate and glutamate methylation in proteins. *J. Pro. Res*., 2008, 7, 1001-6.

[95] Ohvo-Rekilä, H.; Ramstedt, B.; Leppimäki, P.; Slotte, J.P. Cholesterol interactions with phospholipids in membranes. *Pro. Lipid Res*., 2002, 41(1), 66-97.

[96] Lyubartsev, A.P.; Rabinovich, A.L. Recent development in computer simulations of lipid bilayers. *Soft Matter*. 2011, 7, 25-39.

[97] Vattulainen I.; Rog T. Lipid simulations: a perspective on lipids in action. *Cold Spring Harbor Persp. Biol*., 2011, 4, a004655.

[98] Sharma, S.; Kim, B.N.; Stansfeld, P.J.; Sansom, M.S.; Lindau, M. A coarse grained model for a lipid membrane with physiological composition and leaflet asymmetry. *PloS one*. 2015, 10(12), e0144814.

[99] Bennett, W.F.D.; Tieleman, D.P. Computer simulations of lipid membrane domains. Biochim. Biophys. *Acta – Biomemb*., 2013, 1828, 1765-1776.

[100] Accardi, A.; Walden, M.; Nguitragool, W.; Jayaram, H.; Williams, C.; Miller, C. Separate ion pathways in a Cl−/H+ exchanger. *The J. Gen. Physio*., 2005, 126(6), 563-570.

[101] Elvington, S.M.; Liu, C.W.; Maduke, M.C. Substrate-driven conformational changes in ClC-ec1 observed by fluorine NMR. *The EMBO J*., 2009, 28(20), 3090-3102.

[102] Brooks, B.R.; Brooks, C.L.; MacKerell Jr, A.D.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner,

A.R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.W.; Post, C.B.; Pu, J.Z.; Schaefer, M.; Tidor, B.; Venable, R.M.; Woodcock, H.L.; Wu, X.; Yang, W.; York, D.M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comp. Chem*., 2009, 30, 1545-614.

[103] Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comp. Chem*., 2005, 26, 1781-802.

[104] Cheng, X.; Jo, S.; Marassi, F.M.; Im, W. NMR-based simulation studies of Pf1 coat protein in explicit membranes. *Biophys. J*., 2013, 105, 691-8.

[105] Javanainen, M. Universal method for embedding proteins into complex lipid bilayers for molecular dynamics simulations. *J. Chem. Theory Comp*., 2014, 10, 2577-82.

[106] Pruitt, M.M.; Lamm, M.H.; Coffman, C.R. Molecular dynamics simulations on the Tre1 G protein-coupled receptor: exploring the role of the arginine of the NRY motif in Tre1 structure. *BMC Struc. Biol*., 2013, 13,15.

[107] Leontiadou, H.; Mark, A.E.; Marrink, S.J. Ion transport across transmembrane pores. *Biophys. J*., 2007, 92, 4209-15.

[108] Loverde, S.M. Molecular simulation of the transport of drugs across model membranes. *The J. Phys. Chem. Lett*., 2014, 5, 1659-65.

[109] Vaccaro, L.; Koronakis, V.; Sansom, M.S.P. Flexibility in a drug transport accessory protein: molecular dynamics simulations of MexA. *Biophy. J*., 2006, 91, 558-64.

[110] Gumbart, J.; Aksimentiev, A.; Tajkhorshid, E.; Wang, Y.; Schulten, K. Molecular dynamics simulations of proteins in lipid bilayers. *Curr. Opin. Struc. Bio*., 2005, 15, 423-431.

[111] Shimamura, T.; Weyand, S.; Beckstein, O.; Rutherford, N.G.; Hadden, J.M.; Sharples, D.; Sansom, M.S.; Iwata, S.; Henderson, P.J.; Cameron, A.D. Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science*. 2010, 328, 470-3.

[112] Lu, S.Y.; Jiang, Y.J.; Lv, J.; Wu, T.X.; Yu, Q.S.; Zhu, W.L. Molecular docking and molecular dynamics simulation studies of GPR40 receptor–agonist interactions. *J. Mol. Graphics Mod*., 2010, 28(8), 766-774.

[113] Rivaila, L.; Chipotb, C.; Maigretb, B.; Bestela, I.; Sicsica, S.; Tarekb, M. Large-scale molecular dynamics of a G protein-coupled receptor, the human 5-HT 4 serotonin receptor, in a lipid bilayer. *J. Mol. Struc: THEOCHEM*., 2007, 817, 19-26.

[114] Islam, S.M.; Qi, G.F.Y.; Shen, R.; Han, W.; Stein, R.A.; Schulten, K.; Mchaourab, H.S.; Im, W.; Roux, B. Structural-Function Study of Membrane Proteins with Restrained-Ensemble and Dummy Spin-Label Molecular Dynamics Simulations. *Biophys. J*., 2016, 110, 644a.

[115] Feng, T.; Lia, M.; Zhoua, J.; Zhuang, H.; Chene, F.; Yef, R.; Campanellab, O.; Fangg, Z. Application of molecular dynamics simulation in food carbohydrate research-a review. *Inno Food Sci. Emerg. Tech*., 2015, 31, 1-3.

[116] Woods, R.J.; Pathiaseril, A.; Wormald, M.R.; Edge, C.J.; Dwek, R.A. The high degree of internal flexibility observed for an oligomannose oligosaccharide does not alter the overall topology of the molecule. *Europ. J. Biochem*., 1998, 258, 372-86.

[117] Cheong, Y.; Shim, G.; Kang, D.; Kim, Y. Carbohydrate binding specificity of pea lectin studied by NMR spectroscopy and molecular dynamics simulations. *J. Mol. Struc*., 1999, 475, 219-32.

[118] Bernardia, A.; Galganoa, M.; Belvisia, L.; Colombo, G. Simulation of carbohydrate-protein interactions: Computer-aided design of a second generation GM1 mimic. *J. Computer-aided Mole. Des.*, 2001, 15, 117-128.

[119] Ebbinghaus, S.; Meister, K.; Born, B.; DeVries, A.L.; Gruebele, M.; Havenith, M. Anti-freeze Glycoprotein Activity Correlates with Long-Range Protein− Water Dynamics. *J. Am. Chem. Soc*., 2010, 132(35), 12210-12211.

[120] Sapay, N.; Nurisso, A.; Imberty, A. Simulation of carbohydrates, from molecular docking to dynamics in water. *Methods Mol Biol*., 2013, 924, 469-83.

[121] Venkateshwari, S.; Veluraja, K. Molecular modelling and molecular dynamics studies of GD1A, GD1B and their complexes with BoNT/B–perspectives in interaction and specificity. *J. Struc. Biol*., 2012, 180, 497-508.

[122] Grigera, J.R. Conformation of polyols in water. Molecular-dynamics simulation of mannitol and sorbitol. *J. Chem. Soc. Faraday Trans. 1: Phys. Chem. in Cond. Phas*., 1988, 84(8), 2603-2608.

[123] Brady, J.W. Molecular dynamics simulations of α-D-glucose in aqueous solution. *J. Americ. Chem. Soc*., 1989, 111, 5155-65.

[124] Van Eijck, B.P.; Kroon, J. Molecular-dynamics simulations of β-d-ribose and β-d-deoxyribose solutions. *J. Mol. Str*., 1989, 195, 133-146.

[125] van Gunsteren, W.F.; Dolenc, J.; Mark, A.E. Molecular simulation as an aid to experimentalists. *Curr. Opinion Str. Bio*., 2008, 18(2), 149-153.

[126] Mandal, T.K.; Mukhopadhyay, C. Effect of glycosylation on structure and dynamics of MHC class I glycoprotein: a molecular dynamics study. *Biopol*., 2001, 59, 11-23.

[127] Mutter, S.T.; Zielinski, F.; Cheeseman, J.R.; Johannessen, C.; Popelierb, P.L.A.; Blanch, E.W. Conformational dynamics of carbohydrates: Raman optical activity of D-glucuronic acid and N-acetyl-D-glucosamine using a combined molecular dynamics and quantum chemical approach. *Phys. Chem. Chem. Phys.*, 2015, 17, 6016-27.

[128] Wang, D.; Ámundadóttir, M.L.; Gunsteren, W.F.V.; Hünenberger, P.H. Intramolecular hydrogen-bonding in aqueous carbohydrates as a cause or consequence of conformational preferences: a molecular dynamics study of cellobiose stereoisomers. *Eur. Biophys. J*., 2013, 42, 521-37.

[129] Nishiyama, Y.; Johnson, G.P.; French, A.D.; Forsyth, V.T.; Langan, P. Neutron crystallography, molecular dynamics, and quantum mechanics studies of the nature of hydrogen bonding in cellulose Iβ. *Biomacromol*., 2008, 9(11), 3133-3140.

[130] McDowell, S.E.; Špačková, N.; Šponer, J.; Walter, N.G. Molecular dynamics simulations of RNA: an *in silico* single molecule approach. *Biopol*., 2007, 85, 169-84.

[131] Lindert, S.; Bucher, D.; Eastman, P.; Pande, V.; McCammon, J.A. Accelerated molecular dynamics simulations with the AMOEBA polarizable force field on graphics processing units. *J. Chem. Theory Comp*., 2013, 9, 4684-91.

[132] MacKerell Jr, A.D.; Nilsson, L. Molecular dynamics simulations of nucleic acid–protein complexes. *Curr. Opin. Struc. Biol*., 2008, 18, 194-9.

[133] Cheatham, T.E.; Kollman, P.A. Molecular dynamics simulations highlight the structural differences among DNA: DNA, RNA: RNA, and DNA: RNA hybrid duplexes. *J. Amer. Chem. Soc*., 1997, 119, 4805-25.

[134] Schneider, C.; Brand, M.; Sühnel, J. Molecular dynamics simulation reveals conformational switching of water-mediated uracil-cytosine base-pairs in an RNA duplex. *J. Mol. Biol*., 2001, 305, 659-667.

[135] Réblová, K.; Spacková, N.; Sponer, J.E.; Koca, J.; Sponer, J. Molecular dynamics simulations of RNA kissing–loop motifs reveal structural dynamics and formation of cation-binding pockets. *Nucl. Acids Res.*, 2003, 31, 6942-52.

[136] Chen, A.A.; García, A.E. High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Pro. Nat. Acad. Sci*., 2013, 110,16820-5.

[137] Huang, N.; MacKerell Jr, A.D. Specificity in protein–DNA interactions: energetic recognition by the (cytosine-C5)-methyltransferase from HhaI. *J. Mol. Biol*., 2005, 345, 265-74.

[138] Qi, Y.; Spong, M.C.; Nam, K.; Banerjee, A.; Jiralerspong, S.; Karplus, M.; Verdine, G.L. Encounter and extrusion of an intrahelical lesion by a DNA repair enzyme. *Nature*. 2009, 462, 762-6.

[139] Freddolino, P.L.; Harrison, C.B.; Liu, Y.; Schulten, K. Challenges in protein-folding simulations. *Nat. Phys*., 2010, 6, 751-8.

[140] Buck, M.; Bouguet-Bonnet, S.; Pastor, R.W.; MacKerell, Jr, A.D. Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophys. J*., 2006, 90, L36-L38.

[141] Jiang, F.; Han, W.; Wu, Y.D. Influence of side chain conformations on local conformational features of amino acids and implication for force field development. *The J. Phys. Chem. B*., 2010, 114,5840-50.

[142] Skinner, J.J.; Yu, W.; Gichana, E.K.; Baxa, M.C.; Hinshaw, J.R.; Freed, K.F.; Sosnick, T.R. Benchmarking all-atom simulations using hydrogen exchange. *PNAS*. 2014, 111, 15975-80.

[143] Faustino, I.; Pérez, A.; Orozco, M. Toward a consensus view of duplex RNA flexibility. *Biophys. J*., 2010, 99, 1876-85.

[144] Petrov, D.; Zagrovic, B. Are current atomistic force fields accurate enough to study proteins in crowded environments? *PLoS Comp. Bio*., 2014, 10, e1003638.

[145] Yu, I.; Mori, T.; Ando, T.; Harada, R.; Jung, J.; Sugita, Y.; Feig, M. Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *eLife*. 2016, 5, e19274.

[146] Mandell, D.J.; Chorny, I.; Groban, E.S.; Wong, S.E.; Levine, E.; Rapp, C.S.; Jacobson, M.P. Strengths of hydrogen bonds involving phosphorylated amino acid side chains. *J Amer. Chem. Soc.*, 2007, 129, 820-7.

[147] Piana, S.; Donchev, A.G.; Robustelli, P.; Shaw, D.E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *The J. Phys. Chem. B.*, 2015, 119, 5113-23.

Figure legends

Fig. 1. Comparative account of the conventional experimental biology techniques using molecular dynamics.

Before a biological experiment, the model organism needs to be first acclimatized with the lab conditions. The molecular expression of gene, cloning and purification of the gene product and the proteomic studies involve protein isolation, purification and enzyme assays to assess the activity of the protein. The structure related studies are further complicated by the steps involved in protein purification and crystallization. The MD analysis saves the time to be otherwise spent for multiple random trials required for checking the efficiency of mutated proteins and for understanding the possible mechanisms operating in such reactions.

Fig. 2. Readily used different parameters need modifications in the current FFs.

Biomolecules are consists of same kind of atoms but the bond formation, bond angles and their geometry and spatial distribution varies. (A) Different parameters are defined in the FFs in terms of mathematical expression i.e. bond angle, bond length, bonded interactions and non-bonded interactions in the amino acids of the protein. (B) Additional parameters for the torsion angle (χ_1, χ_2, χ_3, χ_4) of the side chain atoms that give rise to torsion parameters of the protein need to be modified. The parameters for elucidating this aspect of the proteins are still under development. (C) The parameters used for nucleic acids are defined on the bond angles of the backbone (α, β, γ, δ, ε, ζ). Since the structure of the nucleic acid is in extended form, however, the complete parameters for explaining the same are not yet well defined. (D) Structure of carbohydrates is

extended and dynamic in the solution. It is not well defined as that of the proteins, which is defined by Ramachandran plot. FFs for carbohydrates are still need to be more developed.

Fig. 3. Flowchart for carrying out simple MD simulation experiments.

The quality of the biomolecular structure needs to be checked before starting the simulation experiment. The functional unit of the biomolecule (monomer, dimer or oligomeric forms) needs to be known. Then the input files are prepared and the biomolecule is immersed into the simulation box. The biomolecule is energy minimized to obtain its most stable computed conformation. The energy minimized biomolecule is equilibrated under constant temperature and pressure conditions. Then it is simulated to study its structural behavior and conformational changes.

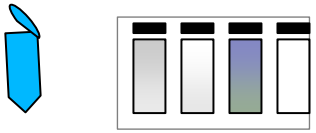
Fig. 4. Applications of molecular dynamics in biomolecular sciences.

MD approaches have been used to elucidate the structure, function and behaviour of biomolecules in duplicate conditions in the simulation box. Different functions of the proteins have been studied by MD analysis (catalysis, drug-binding locations, transport, motility, signal transduction, allosteric regulation, protein-protein interaction specificity/binding, protein embedded in lipids, protein folding/unfolding, post-translational changes, protein-secondary small molecule interactions and molecular recognition of the proteins). The role of carbohydrates involved in the signalling of various cellular processes, hydration of carbohydrates, carbohydrate-protein (glycoprotein) interactions and carbohydrate-ligand interactions are explained by MD. The conformational changes occurring at the molecular levels in DNA/RNA have been successfully studied by MD.

Fig.1.

Plant/Animal

Plant cell/Animal cell

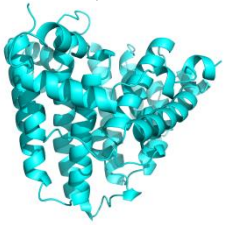


Molecular gene expression studies(DNA/RNA extraction, cloning and gene expression)

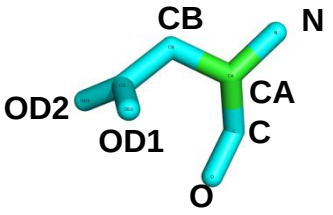
Protein isolation and purification

In vivo/vitro (Protein activities)

MHDLLQTWKADIDGKLKKYQNSVDTNKSSLEKQIA
KQKIFSKQDNIDLMVLNYVRVQAYQSVHEKISLIF



Initial structure of molecule

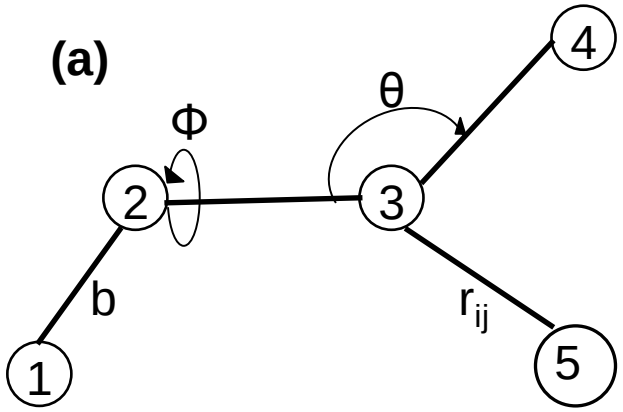


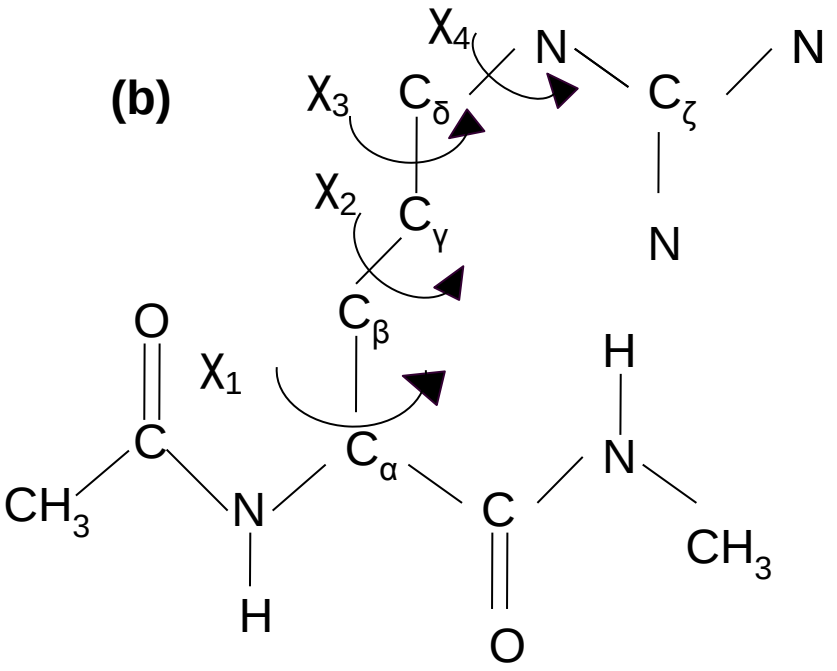
Calculation of molecular forces on each atom

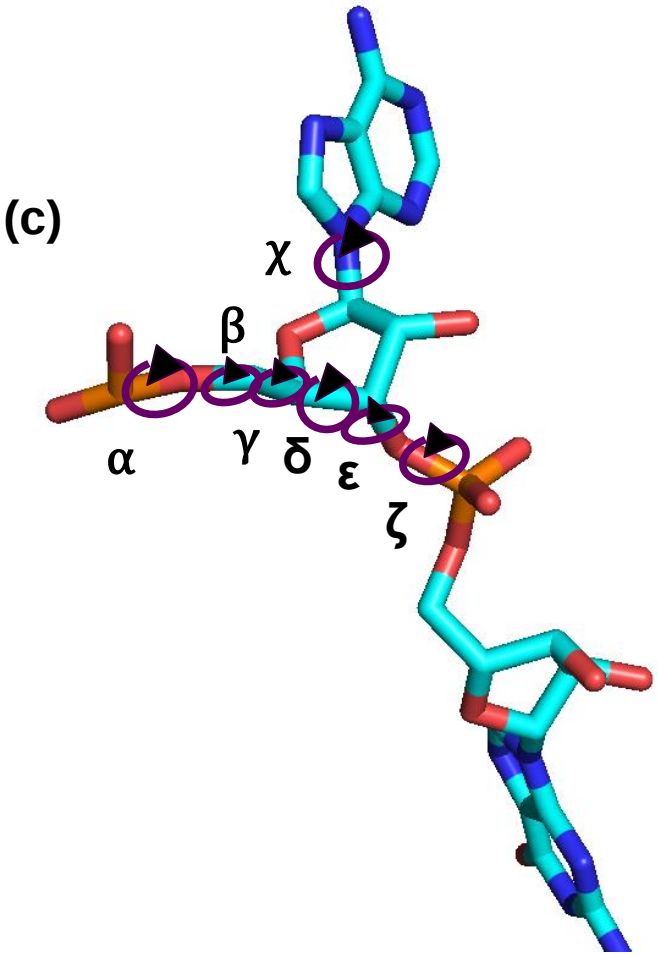
Move each atom according to the force field applied

Promote simulation time by 1/2 fs

Fig.2.


(a)
Φ
θ
2
3
4
1
5
b
r_{ij}


(b)
χ_4
N
N
χ_3
C_δ
C_ζ
χ_2
C_γ
N
C_β
O
χ_1
H
C
C_α
N
CH_3
N
C
CH_3
H
O


(c)
χ
β
α
γ
δ
ε
ζ

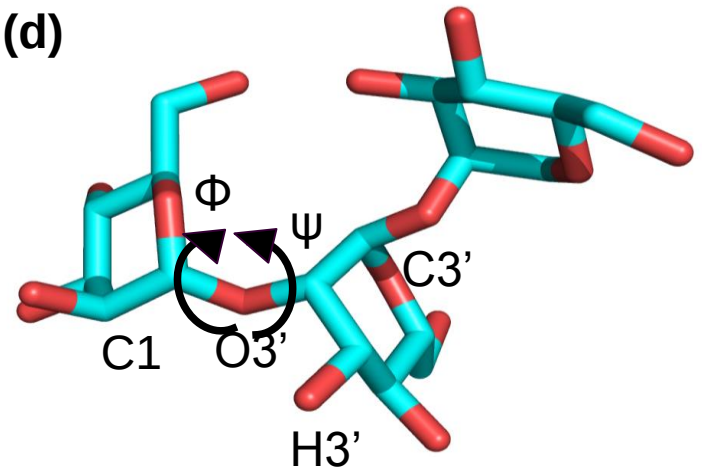

(d)
Φ
ψ
C3’
C1
O3’
H3’

Fig.3.

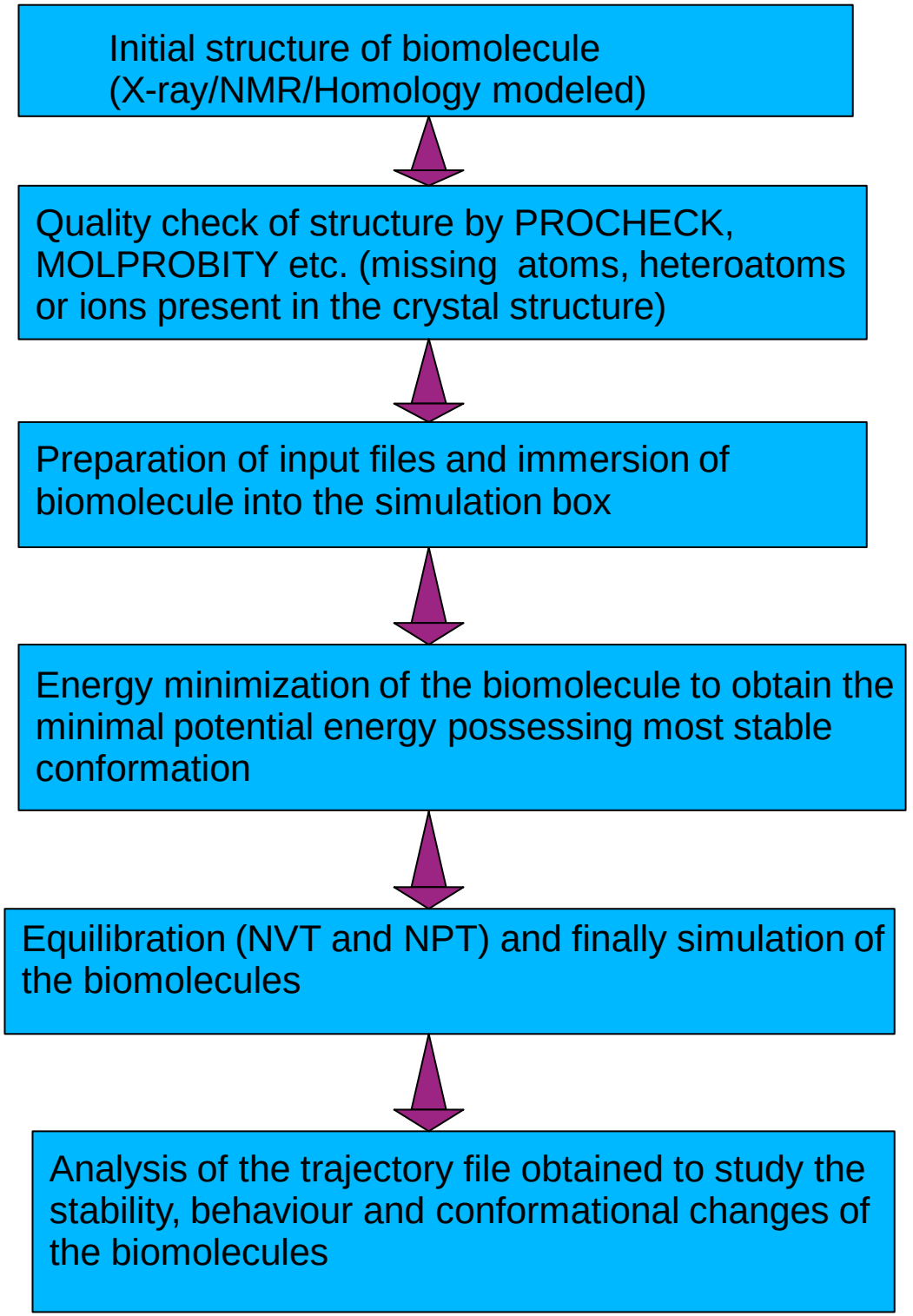

Initial structure of biomolecule (X-ray/NMR/Homology modeled)
Quality check of structure by PROCHECK, MOLPROBITY etc. (missing atoms, heteroatoms or ions present in the crystal structure)
Preparation of input files and immersion of biomolecule into the simulation box
Energy minimization of the biomolecule to obtain the minimal potential energy possessing most stable conformation
Equilibration (NVT and NPT) and finally simulation of the biomolecules
Analysis of the trajectory file obtained to study the stability, behaviour and conformational changes of the biomolecules

Fig.4.