Journal of Molecular Graphics and Modelling 84 (2018) 43–53

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Molecular dynamics simulation for rational protein engineering: Present and future prospectus

Maryam Rouhani [a], Farnaz Khodabakhsh [b], Dariush Norouzian [a], Reza Ahangari Cohan [a, **], Vahideh Valizadeh [a, *]

[a] Nano-Biotechnology Department, New Technologies Research Group, Pasteur Institute of Iran, Tehran, Iran
[b] Department of Genetics and Advanced Medical Technology, Medical Biotechnology Research Center, Faculty of Medicine, AJA University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Article history:
Received 17 April 2018
Received in revised form
5 June 2018
Accepted 8 June 2018
Available online 9 June 2018

Keywords:
Protein engineering
Computational modeling
Glycoengineering
PASylation
PEGylation
Protein stability
Contents

ABSTRACT

Recently protein engineering has been used as a pivotal tool for designing proteins with improved characteristics. While the experimental methods might be laborious and time-consuming, *in silico* protein design is a time and cost-effective approach. Moreover, in some cases, protein modeling might be the only way to obtain structural information where the experimental techniques are inapplicable. Molecular dynamics (MD) simulation is a method that allows the motion of protein to be simulated in defined conditions on the basis of classical molecular dynamics. MD simulation could widely be used when protein design needs accurate modeling of the target protein dynamics and also descriptions of the relation between conformational changes and function of protein at the atomic level. In this review, the effectiveness and the power of MD simulation in designing proteins with improved characteristics will be discussed.

1. Introduction

Proteins are one of the fundamental building blocks of living organisms which are able to form a distinct structure through spatial arrangement. The molecular evolution has revealed that variations occurred in protein sequences, by mutagenesis or recombination, can alter proteins characteristics which is due to generation of new structures [1]. An impressive challenge in structural biology is to design and engineer proteins in order to exhibit new or desired functionalities. Protein engineering is a technology through which novel proteins with preferred or improved properties can be developed. It is one of the dynamically developing disciplines which can be used in bio-industries. In last three decades, the protein engineers have successfully tailored wide ranges of proteins specified to use in industry and medicine [2]. This could be achieved by developing novel experimental technologies such as recombinant DNA, high-throughput screening, deep sequencing, directed evolution, fluorescence-based screening, and gene synthesis [3]. In addition to experimental methodology and rational design strategies, computational methods have been successfully employed for more facile designing and engineering the proteins [4].

2. Computation in molecular biology

Nowadays, computer is a crucial device in most studies, particularly in molecular biology. Computational modeling principally is performed by two different methods; (i) a subjective computer graphic and (ii) an objective computational analysis on the basis of mathematical equations and biophysical properties of structural energies. Translocation, rotation, zooming in real space, creating complicated molecular models, simulation of molecules at atomic levels, analysis of molecular surface, least-squares superposition of molecules, and analysis of large datasets can be accomplished only by computers. Computational methods are mainly grouped into three categories: (i) bioinformatics analysis of primary sequences, (ii) computer modeling of tertiary structures,

* Corresponding author.
** Corresponding author.
E-mail addresses: cohan_r@pasteur.ac.ir (R.A. Cohan), v_valizadeh@pasteur.ac.ir (V. Valizadeh).

https://doi.org/10.1016/j.jmgm.2018.06.009
1093-3263/

known as molecular modeling, and (iii) prediction of new structures by *de novo* design [5].

Molecular modeling studies are often combined with several *in silico* methods such as bioinformatics analysis and quantitative structure/activity relationship (QSAR) to predict effectiveness of every change(s) made in the protein. However, prediction of the conformational behavior of amino acid residues requires accurate estimation of binding energies and assessing the reaction activation barrier changes. In spite of significant progress in quantum mechanics, in the long-term simulation of complex molecules, MD simulation is still preferred due to the greatly reduced computing process [3,6]. MD simulation is a powerful method to study the dynamic feature of a protein at atomic levels [7]. Additionally, it provides correlation between structure and dynamics considering conformational energy landscape accessible to protein molecules [8–10]. Today, modern computers allow molecular simulations ranging from nanoseconds to microseconds which are enough time to determine conformational changes at atomic level. Therefore, MD simulation is an attractive method which identifies flexible regions in the protein serving as proper targets for stability increment or achievements in other protein engineering objectives [7].

3. Molecular dynamics simulation for protein engineering

Earlier, protein modeling software tasks dealt with chain closure, constructing molecules from building elements, and examining the conformational space by manual changes of torsional angles [5]. In 1972, Katz and Levinthal studied the hardware and software aspects of molecular structure presentation, manipulation, and structure fitting into electron density contours. Beside many aspects of protein structural science, computer hardware progress and evolution of methods revealed direct impact in this field. For instance, development of computer graphic was often driven by technical requirements of protein crystallography techniques [11].

One of the early papers published on computer modeling of protein was the study of Levitt and Marshal (1975). They described the computer modeling of a protein folding on the basis of a new and simple presentation of a protein conformation plus energy minimization and thermalization. The method successfully described the renaturing of bovine pancreatic trypsin inhibitor from open-chain conformation to the folded state which is similar to the native protein under defined conditions [12]. Later, Sherga and Kuntz published their studies with similar themes [13–15]. However, in a primitive modeling system, despite the significant simplification of the protein molecule, the molecular mechanic force fields were not selected properly. Because the potential energy of protein in a vacuum is not a good approximation for the free energy of a biological system [5]. Computer modeling should finally face the biophysical and thermodynamic characteristics of a protein in an aqueous solution. A decade later, a 210-ps simulation of understudied protein in water was reported [16].

Later, a significant increase in computing power brought about routinely simulation of larger proteins which are 1000–10000 times longer than the original primitive simulation in an aqueous environment containing ions. Additionally, significant improvements in the potential function of a protein with respect to enhancing the stability have been achieved [17]. This potential was attained by using more defined/accurate force fields along with coordinates [18]. In theory, the force field is employed for parameterizing protein energy. However, considering the structural complexity of protein, protein force fields are divided into different terms. Regarding variety of methods existing to develop a model system and parameterize surface energy of a protein, different force fields are available [19]. Currently, the most commonly used force fields are chemistry at HARvard molecular mechanics (CHARMM) [20], assisted model building and energy refinement (AMBER) [21], Optimized potential for liquid simulations (OPLS) [22] and groningen molecular simulation (GROMOS) [23]. These force fields are generally available in particular modeling packages which frequently can be used to simulate the macromolecules.

The ultimate goal of protein modeling is the accurate prediction of a protein structure from its primary sequence which is comparable to that of experimentally obtained results [24]. This will allow the investigators to safely use easy generated *in silico* protein models which can be used in all contexts, instead of experimental examination. Such approach could be very supportive in structure-based drug design, analysis of protein function, rational design of proteins with enhanced stability or increased *in vivo* half-life, decrement in immunogenicity, and in some cases, achieving novel proteins with new functionalities [25–27]. *In silico* design is an alternative time saving and cost effective method which sometimes is the only approach to obtain the structural information of the protein when the experimental procedures are failed [24]. Regarding this viewpoint, in this manuscript, the attempts will be made to review the application of MD simulation in protein engineering. In the following, a number of protein modification approaches will be discussed in which molecular dynamics was used for simulating engineered proteins that led to a general understanding of mechanism or developing a molecule with enhanced or novel properties.

3.1. Molecular dynamics simulation for protein glycoengineering

A large number of therapeutic proteins have been developed for the treatment of different diseases, but some drawbacks, like loss of activity or rapid clearance from the circulation, limit their clinical applications [28]. Novel strategies are in use to design new drugs with higher activity and longer *in vivo* half-life. Glycoengineering which means a change in carbohydrate moiety of a protein, causes alterations in pharmacokinetics characteristics of the target protein [29,30]. Carbohydrate chain addition to the protein can lead to a significant reduction in protein aggregation by increasing solubility through masking the hydrophobic patches on the protein surface [31].

Naturally, protein glycosylation involves covalent binding of glycan to proteins through amino acid side chains of asparagine (N-linked), and serine/threonine (O-linked) [32]. The N-glycosidic linkages occur between the carbohydrate moiety of beta-N-acetyl glucosamine and the side chain of asparagine residue wherein the amino acid is embedded in tripeptide sequence N-X-S/T [33,34]. On the contrary, the O-glycosidic linkage occurs between different glycan moieties and the residues of serine/threonine. Unlike the particular position of an asparagine residue in tripeptide sequence, serine/threonine involved in O-linked glycosylation does not show any specific amino acid sequence preferences.

Introduction of new glycosylation sites in the protein structure leads to formation of proteins with high carbohydrate content. Nonetheless, introducing new glycosylation sites affects the folding, three-dimensional (3D) structure, and activity of target protein [29]. Moreover, the surface accessibility of asparagine residue and the possibility of enzymatic glycosylation should be considered. Therefore, rational selection of proper positions for the introduction of new glycosylation site before the experimental approach is important due to being cost-effective and time-saving.

Samoudi et al. (2015) used recombinant human β interferon (rhuIFN-β) as a model protein to identify the suitable positions for introducing new N-glycosylation sites [35]. They employed a computational strategy to predict the structural distortion and function of the target protein which might be caused by the

changes in amino acid sequence. In this way, the 3D structure of rhIFN-β analogues was designed by comparative modeling. Then, MD simulation was employed to evaluate the molecular stability and flexibility profile of the predicted structures. In addition, accessibility of asparagine residues in the novel generated structures was investigated. This computation strategy could be applied to avoid changes in the 3D structure of protein caused by alteration in amino acid sequence while designing the novel glycosylated forms (Fig. 1).

The experimental evaluation of the two suggested analogues (L6T, S57 N) indicated that the new sites were properly glycosylated and the resultant glycoengineered rhIFN-β forms retained their functions, thus revealing the validity of MD simulation methodology (Table 1). Further analysis confirmed the accuracy of such computation with an improvement in the aggregation state of the engineered proteins [36]. This kind of computational study for predicting and selecting the hyperglycosylated analogues was of great importance, due to the complexity of construction, expression, and characterization of different analogues which is time-consuming and labor-intensive. For instance, Elliot and colleagues constructed numerous human erythropoietin analogues (62 analogues) and studied experimentally their new sites of hyperglycosylation [37].

In 2016, Ghasemi and coworkers employed a similar method to design hyper-glycosylated forms of human coagulation factor IX. In their research, proper new glycosylation sites in the protein were defined by MD simulation and five hyper-glycoengineered forms were selected. The data obtained during MD simulation resulted in one analogue showing similarity to structure and dynamic/kinetic behavior of the native protein which was taken into future studies [38].

In another study, the impact of N-glycosylation on the structure and dynamics of protein was investigated systematically the integrated computational method of Protein Data Bank (PDB) structure analysis and atomistic MD simulation of glycosylated and non-glycosylated proteins. In this investigation, the effect of N-glycosylation on overall and local structure of protein was determined by modeling the structural similarity between glycosylated and deglycosylated forms of identical proteins in PDB. Further, the overall structural similarity was quantified by determining the root mean square deviation (RMSD) and the TM score. The results indicated that N-glycosylation did not induce significant conformational changes; however, RMSF analysis of MD trajectories from 6 representative glycoproteins and their de-glycosylated forms revealed that N-glycosylation decreased the dynamic fluctuation of protein, probably leading to protein stability [39].

As it has already been stated in some cases, protein modification experimentally is hard to be practical and thus computational protein modeling could be the method of choice. In this regard, attachment of N-linked glycans to specific amino acids of gp120 envelope trimer of an HIV virion can modulate the binding affinity of gp120 to CD4, influencing co-receptor tropism and plays an important role in neutralizing antibody responses. Due to challenges associated with crystallization of glycosylated proteins of HIV virion, most of the structural studies are focused on describing the features of a non-glycosylated HIV-I gp120 protein. Wood et al. employed a computational approach to determine the effect of N-linked glycan on dynamics of gp120 protein HIV, particularly v3 loop.

In this study, the MD simulation was carried out to compare the structured form of non-glycosylated gp120 with two glycosylated gp120 structures. One of the glycan forms possesses a single covalently linked high-mannose glycan and the second form contains five of this glycan moiety. They used the tLEaP module of AMBER 10 [40] with the AMBER ff99SB [41] force field for the protein, and GLYCAM06 [42] force field for the carbohydrates, to generate the coordinate and topology files for both the glycosylated and the non-glycosylated proteins. Their results illustrated a significant effect of N-linked glycosylation on sequential and 3D properties of the underlying protein structure. They showed that glycan surrounding v3 loop regulates its dynamics and conferring more narrowly conformational changes to the loop as compared to its non-glycosylated counterpart. The conformational effect on v3 loop approves that N-linked glycosylation plays an important role in determining HIV-I co-receptor tropism [43].

3.2. Molecular dynamics simulation for protein PEGylation

Polyethylene glycol (PEG) is a polymer with the ability of sterically shielding molecules (pharmaceutical or biopharmaceutical) either through encapsulation or covalently/non-covalently bound to the drug vehicles in a process which is known as PEGylation [25]. It was introduced for the first time by Abuchowski in 1977 [44,45]. PEGylation was accepted as a rapid, most effective and widely applied approach to circumvent cytotoxicity and increment in solubility and circulation time of the drug molecules or drug carriers [46–48]. Though fundamental information about the binding and large scale interaction of PEG in drug molecules or drug carriers can be obtained through experiments, but circumstances occur during this binding in atomic level, cannot be explained by

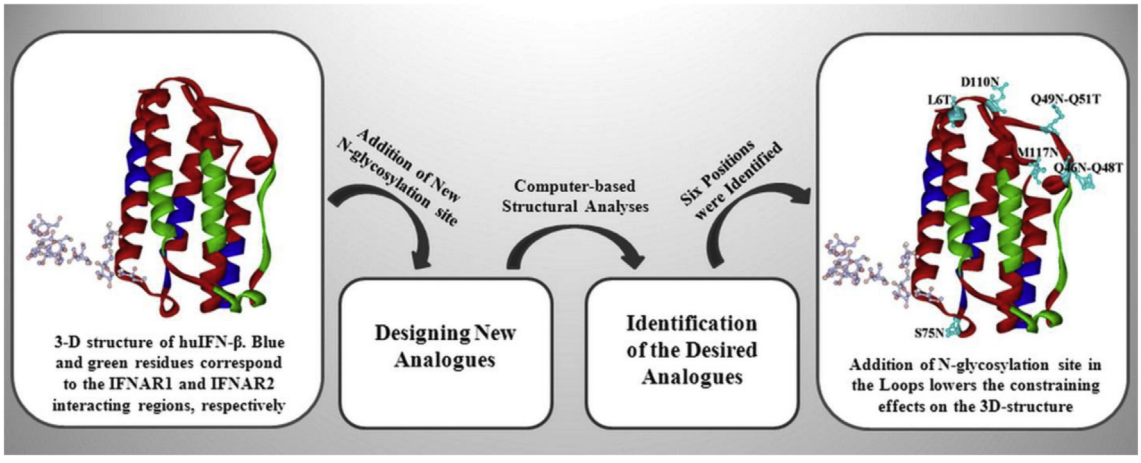

Fig. 1. Schematic representation of interferon beta analogues for glycoengineering. Reprinted with permission from Ref. [34].

Table 1
Evaluation of human Interferon beta analogues for glycoengineering based on whole-structure RMSD, flexibility, surface area accessibility after MD simulation.

Whole-structure RMSD relative to the wild-type (nm)	ASA value for Asn residue in new glycosylation site (%)	Consistency of flexibility profile	Average RMSDs over the equilibration phase (nm)	Analog
0.20	53 ± 6	Yes	0.27 ± 0.011	L6T
N/A	N/A	No	0.26 ± 0.012	R11 N
N/A	N/A	No	0.25 ± 0.022	A56 N
0.19	93 ± 8	Yes	0.25 ± 0.012	S75 N
N/A	N/A	N/A	N/A (unstable)	L98T
0.17	0	Yes	0.24 ± 0.01	L98 N
0.16	91 ± 14	Yes	0.26 ± 0.011	D110 N
0.19	86 ± 4.5	Yes	0.3 ± 0.013	M117 N
N/A	N/A	N/A	N/A (unstable)	R159 N
0.15	80 ± 9	Yes	0.25 ± 0.013	Q46N-Q48T
N/A	N/A	No	0.3 ± 0.009	Q48N-F50T
0.18	84 ± 7	Yes	0.26 ± 0.006	Q49N-Q51T
N/A	N/A	N/A	N/A (unstable)	Q51N-E53T
0.20	9 ± 7	Yes	0.26 ± 0.011	L6N-F8T
N/A	N/A	N/A	N/A (unstable)	D54N-A56T
N/A	N/A	N/A	N/A (unstable)	K115N-M117T

Reprinted with permission from reference (34).

experiments and remain unanswered. Thereby molecular dynamics simulation can be employed to resolve such interaction at atomic level [49]. Further, conventional methods used for PEGylation led to the formation of heterogeneous mixture and low activity profile of the target molecules. Currently the attempts are made to employ site-specific approaches to PEGylate the target protein via chemoselective targeting of amino acids [50]. Computational analysis is highly recommended to select the proper position on the protein for PEGylation.

One of the studies to resolve the process of PEGylation at atomic level was carried out by Yang et al., (2011) [51] employing conjugated protein as represented by insulin initially. Then MD simulation was performed to assess the interaction between insulin and the different chain lengths of PEG polymer at atomic level through GROMACS package version 4.0.5 [52,53] employing GROMOS 96 43a1 force field [54]. The conjugated molecule was solvated by explicit water molecule through simple point change model [55]; accordingly in order to neutralize the system, the ions were added. The simulation was carried out under the triclinic cell geometry and periodic boundary conditions. During the simulation insulin molecule conjugated to different length chains of PEG polymer or unconjugated insulin was placed at the center of a simulation box with appropriate size. The results revealed, PEG entangled around the protein surface through hydrophobic interaction and on the other hand established hydrogen bonds with the surrounding water molecules. Define secondary structure of protein (DSSP) algorithms showed that the secondary structure of PEG-insulin was unmodified during simulation. The structural stability was improved as shown by RMSD and secondary structure analysis. Data analysis showed that PEGylation caused increment in size of insulin while decreasing the solvent accessible surface area of the protein. PEGylation brought about the prolong circulation due to decrease kidney filtration, resistance to proteolysis and reducing immunogenic side effects as demonstrates experimentally [51].

In another study, the evaluation of steric shielding of PEG on the ester groups of the protein under modification was carried out in order to provide more clearly molecular insight into the interaction between the PEGylated protein and its receptor. Mu et al., 2013 [56] investigated four forms of PEGylated staphylokinase (SK) as prepared by site specific conjugation of 5 KD and 20 KD of PEG to the N and C termini of SK, respectively. Structural evaluations suggested that the native conformation of SK that supported its bioactivity essentially did not change upon coupling to PEG polymer. Although, the PEG chain length and the PEGylation site may alter the hydrodynamic volume and molecular symmetry of the examined protein. MD simulation of PEGylated SKs showed that PEG polymer remained flexible at equilibrium state. PEG polymer is capable of forming a distinctive hydrated layer around SK, which protects the enzyme, maintains the native conformation and increases the hydrodynamic volume. Further, the results obtained by simulation suggested that PEG polymer wraps around the protein providing steric shield. The protein and such steric shielding effect depend upon the PEG chain length and site of PEGylation of the protein. Interestingly, the receptor binding domain of SK is less sterically shielded by PEG in C-terminally PEGylated SK, as reflected by the lower solvent accessible surface area (SASA) of its receptor binding domain. It is assumed by the fact that N-terminus is closer to the receptor binding domain than the C-terminus domain. In conclusion, the results obtained experimentally for PEGylated SK agreed with that of computational simulation [56]. This type of study is actually needed for the rational design, fabrication and further clinical application of PEGylated proteins when they involve the interaction with high molecular weight substrates or receptors.

MD simulation can also be applied for simulating the engineered molecules through site specific PEGylation. PEGylation of protein through cysteine content of the protein is one of the most important techniques applied to PEGylate the proteins [57]. In such method, some of the amino acids which are not involved in the activity of the protein are replaced by molecular biology techniques and then 3D structures of the altered proteins were generated using homology modeling. Based on the result obtained by MD simulation, the stable modified proteins with the inserted cysteine residue at the surface accessible in the structure can be selected. In this regard, our research group studied the site specific PEGylation of erythropoietin by inserting a cysteine residue at the surface of target molecule which is accessible for PEGylation. Erythropoietin (EPO) is a glycoprotein hematopoietic growth factor with short *in vivo* half-life which limits its therapeutic usage [58]. PEGylation is one of the most promising tools, performed for *in vivo* half-life extension of EPO.

In this regard, Maleki et al., 2012 [59] generated 3D structures of human cysteine recombinant erythropoietin (Cys-rhEPO

analogues) by homology modeling using MODELLER software. The quality of the models was checked by discrete optimized protein energy (DOPE) score profile and the Ramachandran plot. Furthermore, the stability of modeled Cys-rhEPO was checked by MD simulation by explicit water using GROMACS 3.3 package and GROMOS 96 force field [52]. The average structures corresponding to the last 200–500 ps of MD simulation were calculated and Cys-rhEPO analogues having the RMSD less than 4 Å from the initial structure were selected for further studies. RMSD of all atoms between native and EPO analogues were calculated by Qmol software [60]. The surface area accessibility of engineered cysteine residues in modeled Cys-rhEPO was computed by GETAREA server [61]. In the next step, the cDNAs coding for Cys-rhEPO were constructed by site directed mutagenesis and expressed as heterologous proteins in *Pichia pastoris*. The results suggest that PEGylated EPO analogues expressed in *Pichia* can be considered as a promising approach to generate cost effective and long acting erythropoiesis stimulating agents [59] (Fig. 2).

Furthermore, Cohan et al. studied the chemo-selective PEGylation of rhEPO by molecular dynamics and graphic *in silico* techniques in order to overcome rhEPO pharmacokinetics limitations. The most appropriate analogue was selected through MD simulations and then expressed in Chinese hamster ovary (CHO)/dhfr$^-$ cells (Table 2). In this way the PEGylated rhEPO analogues were obtained selectively. The results indicated a nano sized EPO31-PEG showing five-fold longer half-life in rats with similar biological activity as compared to native rhEPO [62]. In another investigation, cysteine site specific PEGylation of human basic fibroblast growth factor was performed with the similar approach. In this study D68C, K77C, E78C, and R81C analogues were successfully applied for experimental steps based on minimum protein energy and RMSD differences form the native protein [63].

Mirzaei et al., 2016 [64] investigated cysteine PEGylation of erythropoietin analogues. Their examination was based on computational and non-glycosylated systems to define a simpler approach for site specific PEGylation. Thereby two analogues (E31C and E89C) were chosen on the basis of lowest structural deviation from the native form, accessibility, and nucleophilicity of the free thiol group for PEGylation (Fig. 3). The non-glycosylated erythropoietins E31C and E89C were then PEGylated by mPEG-maleimide (20 kDa) whichresulted in 79% and 82% conjugation yield, respectively. The size and charge characteristics showed an increase in size and negative charge for PEGylated forms as compared with the non-PEGylated EPO. As a concluding remark, development of cysteine-PEGylated proteins by using non-glycosylated expression systems and *in silico* approach can be considered as an efficient technique for optimization of PEGylation parameters [64].

Table 2
Evaluation of the stability and surface area accessibility human erythropoietin cysteine analogues using MD simulations for site specific PEGylation.

SAA (%)	Stability	Cysteine analog	SAA (%)	Stability	Cysteine analog
6.1	Stable	E117C	48.8	Stable	A1C
2.8	Stable	A118C		Unstable	P2C
47.4	Stable	I119C	6.7	Stable	P3C
	Unstable	S120C	54.2	Stable	R4C
17.6	Stable	P121C	60.6	Stable	L5C
7.7	Stable	P122C	22.1	Stable	I6C
88.7	Stable	D123C	57.1	Stable	T27C
0.7	Stable	A124C		Unstable	G28C
51.8	Stable	A125C		Unstable	A30C
28.2	Stable	S126C	100.0	Stable	E31C
30.1	Stable	A127C	45.6	Stable	H32C
0.0	Stable	A128C		Unstable	S34C
3.7	Stable	P129C	0.0	Stable	L35C
78.1	Stable	L130C	48.1	Stable	N36C
25.8	Stable	T132C		Unstable	E37C
41.1	Stable	I133C	0.0	Stable	V41C
52.2	Stable	T134C	10.3	Stable	P42C
	Unstable	A135C		Unstable	Q86C
	Unstable	D136C	48.7	Unstable	P87C
43.7	Stable	T137C	0.9	Stable	W88C
72.9	Stable	R162C	96.6	Stable	E89C
51.9	Stable	T163C	9.4	Stable	G113C
42.4	Stable	G164C	0.0	Stable	A114C
9.0	Stable	D165C	46.8	Stable	Q115C
70.8	Stable	R166C	72.6	Stable	K116C

Reprinted with permission from reference (61).

3.3. Molecular dynamics simulation for protein PASylation

Protein PEGylation can frequently be used to improve the bio drugs stability and increasing blood circulation half-life. Though it is theoretically simple and efficient method but there are serious practical and medicinal disadvantages. The important drawbacks include hyper sensitivity, unexpected changes in pharmacokinetic behavior of conjugated protein, toxic side products formation, easily degradation of the coupled protein under mechanical stress and possible accumulation in the body due to its non-biodegradability which can limit its usage [65]. However different polymeric materials have been proposed as an alternative to PEG polymers for protein conjugation, but eventually chain of amino acid sequences, which are intrinsically unstructured, were introduced. The most advantage of these polypeptide chains is high solubility with no effect on charges of coupled protein. Firstly, a

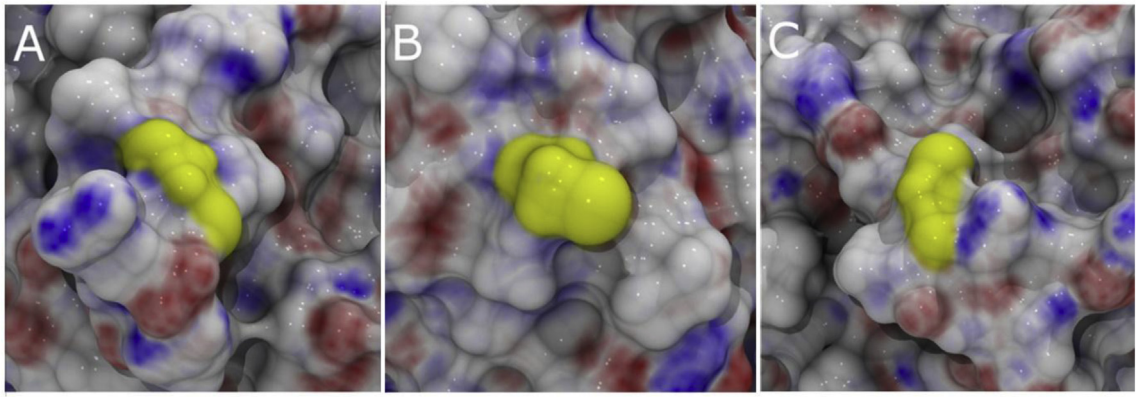

Fig. 2. Conformational investigation of the replaced cysteine residues (yellow color) in the averaged structure of mutants; (A) A1C, (B) E89C, and (C) R162C analogues. Reprinted with permission [58]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

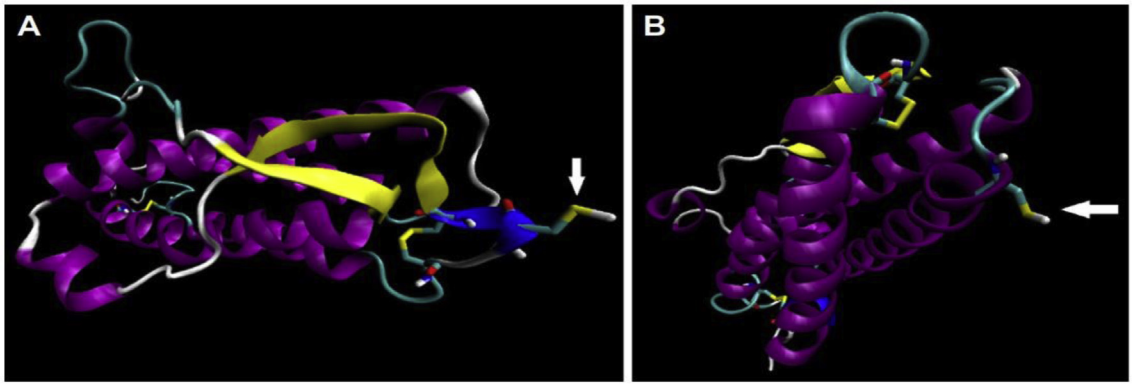

Fig. 3. Ribbon like structure representation of (A) E31C, and (B) E89C non-glycosylated erythropoietin mutants. The white arrows show the free accessible engineered cysteine residues for conjugation with the PEG polymer Reprinted with permission [63].

polypeptide chain with the glycine backbone which exhibiting a random conformational structure, was developed. This type of polypeptide imposed large, hydrodynamic volume on the target protein [66]. Despite achieving significantly prolonged half-life, but due to limited solubility of glycine, preparation of long glycine chains was prevented. As it is known that homopolymer of small amino acids adopt β-sheet and/or helices rather than random coil, combination of different residues were explored. Schlapschy and colleagues found that sequences composed of amino acids like proline, alanine, and serine (PAS) permitted the efficient bio synthesis of long hydrophilic polypeptide chains. Interestingly such highly soluble polypeptide chains demonstrated properties similar to PEG polymers [67].

Pro/Ala rich polypeptide sequence, particularly those containing PAS establish highly hydrophilic biological polymer that can be directly produced as a fusion protein with biopharmaceutical properties through standard expression systems. As compared to PEG, PAS sequences cause increment in the size of conjugated target protein which is larger than the size of renal glomeruli thus mitigating kidney filtration [67]. Polymer of PAS containing 200–600 residues have been demonstrate to increase the plasma half-life for about 10–100 folds [68]. As it is already mentioned, *in silico* studies and evolution of PASylated molecules by MD simulation is effectively beneficial.

Recently, Hedayati et al., 2017 [69] investigated the ability of PASylation technology to improve pharmaceutical properties of rhEPO. They performed modeling prediction as follows: NCBI protein BLAST tool was used for obtaining suitable template structure of EPO, and then 10,000 models were designed with MODELLER software based on obtained template (1EER, 1.9 Å, X-Ray Diffraction). In parallel, I-TASSER server based on threading fold recognition method was used to model PAS sequences. Thereafter, the models for 3D structures of EPO and PAS were utilized as new templates to obtain the full length of EPO-PAS model by MODELLER software. Molecular dynamics of the models were analyzed by MD simulation through GROMACS 4.5.3 package on Red Hat Linux with OPLS-AA force field. Additionally the quality of the generated models (φ and ψ angels) was assessed by RAMPAGE server. The analysis of phi & psi angels demonstrated that more than 99.2% of the residues were located in the allowed regions of Ramachandran plot (Fig. 4). Interestingly, experimental results were in parallel with *in silico* techniques; hence, the molecular weight of purified rPAS-EPO was equal to the expected value. The *in vitro* study showed that the biological activity of fused protein was significantly reduced; however, the *in vivo* activity was considerably increased as indicated by normocythemic mice assay. In addition, pharmacokinetics studies on animal showed highly plasma half-life extension for PASylated EPO [69].

Interestingly, in a very recent published work by our group, the effect of PAS#1 (200) polypeptide on the spatial structure of anti-vascular epidermal growth factor type A (VEGF-A) nanobody was investigated *in silico*. The 3D-structure of PAS-fused nanobody was generated using I-TASSER and Chimera. Energy minimization of the 3D-model was carried out by conjugate gradient method for 50 ns by using nanoscale molecular dynamics (NAMD) package. The visualization and RMSD analysis of the energy minimized structure did not show deleterious effects of PAS sequence on the structure of nanobody. The experimental data have also confirmed that the biological activity of nanobody was not altered after the attachment of PAS#1 (200) sequence [70].

3.4. Molecular dynamics simulation for protein stability

One of the most important aims in protein engineering is the increment of protein stability against temperature, solvent and other harsh conditions. The stability of proteins could be due to several factors such as hydrophobic, electrostatic interactions and also hydrogen bonding. One of the techniques to increase the stability of protein is the introduction of disulfide bond to the protein structure which is attributed to lowered entropy of the unfolded state of the examined protein [71]. Furthermore, introduction of disulfide bonds into an enzyme kinetically stabilized the protein and this is due to the involvement of the disulfide bonds in protein amongst the regions in early stage of the unfolding process [72]. Mansfield et al., 1997 [73] illustrated extremely thermal stability increment of a protease (thermolysin) by introducing a disulfide bond in a region involved in partial unfolding process. Earlier reports suggest the successful introduction of disulfide bonds to the native unpaired cysteine residues which are the indicator for another cysteine residue in protein structure [74,75]. Later Pantoliano et al. demonstrated stabilization of protein can be achieved by *de novo* introduction of paired cysteine residues [76]; hence, the number of theoretically introduced disulfide bonds can be more than the practically possible ones.

Theoretically introduction of disulfide bridges could stabilize the protein but in some proteins the reverse effects have been observed. The destabilization effect might be due to negative effects of a rigid bond construction on packing and flexibility of the target protein. The substantial information for spatial requirements of disulfide bond must be available by high resolution structures [77]. This is brought about by developing computer programs that are capable of predicting proper sites to introduce new disulfide

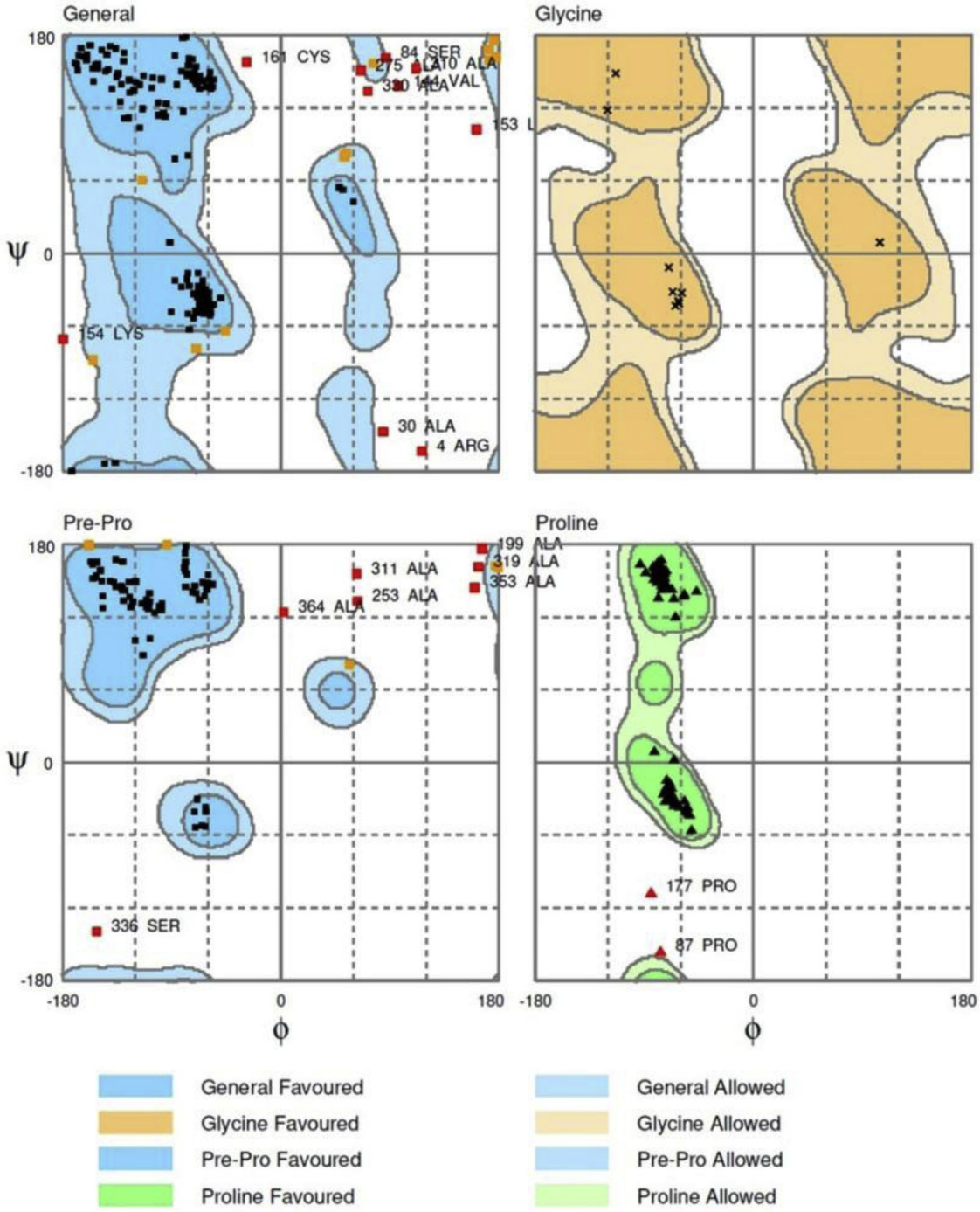

Fig. 4. The Phi and Psi analysis of the final 3D structure of PAS200-erythropoietin. As indicated in the figure, more than 99.2% of residues are located in the allowed regions. Reprinted with permission [68].

bonds [78–80]. Although, such programs facilitated to explore the successful substitution, however, the introduction of disulfide bonds for stabilization of larger proteins is a matter of trial and error. In fact it is not possible to identify the regions in target protein that are essentially important for higher stability. In this regard, determining the regions which are flexible and are the sites for the introduction of disulfide bonds by MD simulation can extremely reduce the trial and error examination.

In this regard, Haloalkane dehalogenase (DhlA) was studied as a model protein in order to identify an appropriate region for introduction of disulfide bridge by MD simulation technique. DhlA of bacterial origin (*Xanthobacter autotrophicus* GJ10 (DhlA)) catalyzes hydrolytic dehalogenation of vast range of small halogenated molecules. The enzyme belongs to α/β-hydrolase family and comprises two domains: i) α/B hydrolase fold domain and ii) cap domain consisting of 5 α-helices [81]. A disulfide bond prediction computer program, SSBOND [79], was employed to choose proper sites for the introduction of a disulfide bond bridging the helix-

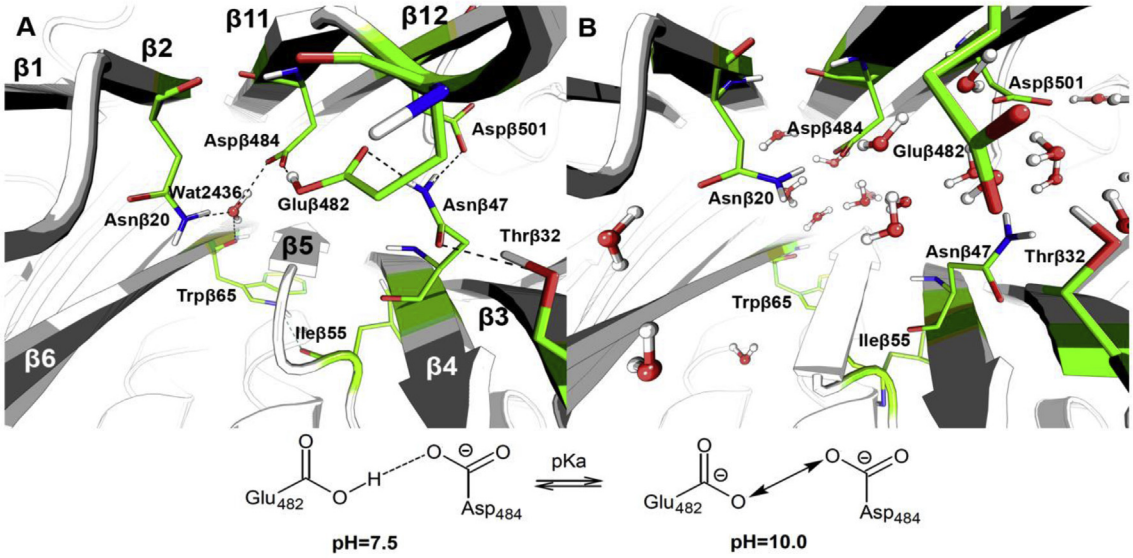

Fig. 5. The hydrogen bonding network between the two β-layers of the αββα-core B1 domain at pH 7.5 (A) and pH 10.0 (B). Reprinted with permission [83].

loop-helix region and the main domain [82]. For this, an X-ray structure was obtained at pH 8.2 having the resolution of 1.9 Å (PDB, IEDE) which served as an initial structure [83]. In continuation, the mutant D16C/A201C was generated and both mutated and native proteins were simulated. MD simulation of Dh1A revealed high mobility in a helix-loop-helix region in the cap domain. The mutated protein demonstrated significant changes in thermal behavior by urea assay. Moreover, the apparent transition temperature of oxidized form of mutated protein was increased by 5° C as compared to its native form. The same trend was observed in urea denaturation assay. Thus the results expressed that identification of mobile domains in 3D structure of protein understudy by MD simulation led to the identification of sites where disulfide bonds could properly be inserted in order to enhance the protein stability [7].

In the other study performed by Liu and Wang (2003), thermo stability of glucoamylase from *Aspergillus awamori* was increased dramatically by introducing six disulfide bonds which 'lock' the α-helix 11 on the surface of the catalytic domain. The catalytic center was well protected by the (α/α)6-barrel at MD simulation temperatures up to 600 K [84]. Moreover Interlandi et al. .2003 showed that Ankyrin Repeat Proteins (ARPs) could be more thermo stable with one to five repeats. In these proteins the stability of the native structure was found to increase with the number of repeats. They carried out multiple MD simulations with explicit solvent at room temperature and at 400 K to characterize designed ankyrin repeat (AR) proteins with full-consensus repeats [85]. In the another interesting study, Suplatov et al., 2014 [86], studied pH stability of penicillin acylase produced by *E. coli* (EcPA) through MD simulation. Their report suggested that denaturation of PA under alkaline conditions could be occurred by transition to an irreversible unfolded state. Thus the conformational mobility of PA was studied by simultaneous MD simulations under neutral and alkaline pH to find out the instability mechanism and the residues critically involved in stability of PA under such conditions. Followings were the approach to achieve pH stable EcPA. Nanosecond MD simulation using NAMD version 2.9 [87] was done at pH 7.5 and 10.0 at 373 K. Then structural differences checked by calculation of RMSD of the backbone in trajectories using visual molecular dynamic (VMD) program [88]. Furthermore, the different conformational mobility of EcPA at various conditions was diagnosed by Δ- RMSD and *f* values and structural flexibility was evaluated with RMSF calculations and analysis of time-evolution secondary structure using Timeline plugin by VMD. Parameterization was performed by Amber tools package version 12 [89]. They concluded that a hydrogen bond exists between Eβ482 and Dβ484 acts as a center of buried stabilizing interaction network at neutral pH. It can be explained by MD simulation, the distance between the side chain oxygen of Eβ482 and Dβ484 is 2.5 Å and thus two carboxyl groups make a hydrogen bond. At neutral pH Eβ482 is protonated to be involved in hydrogen bonding, and at alkaline pH it is deprotonated and becomes negatively charged. Thus, the hydrogen bond network was collapsed due to repulsion between Eβ482 and Dβ484 leading to loss of protein functionality (Fig. 5). Mined data revealed that Dβ484 can be selected as a site for enzyme engineering to tolerate alkaline pH. MD simulation of mutated proteins showed hydrogen bond formed between the mutated residues of Dβ484 N and Eβ482 can maintain the stabilizing interaction network at different pHs understudy. Finally, Dβ484N-EcPA mutant experimentally showed 9 fold increase in activity at pH 10.0 [86].

3.5. Molecular dynamics simulation for protein redesigning

Proteins can be redesigned based on our knowledge about the characteristics of the target protein, permitting the development of novel and unnatural proteins which might be applied in relevant industries. Computational approaches and direct evolution can provide strong and efficient strategy to engineer proteins for achieving desired activity/functionality which is considered as rational design of protein. In the case of enzyme improvement, firstly, residues taking part in substrate binding or transition state should be identified by computational approach and then mutagenesis of the desired site to be carried out by direct evolution through computed guided oligonucleotide directed mutagenesis.

In this regard, sperm whale myoglobin (swMb, PDB: 1JP6) was chosen as a structural and functional model of a metalloprotein, nitric oxide reductase (NOR), to engineer a NOR active site. It was easy to prepare and crystalize swMb and compare it to native and synthetic models of NOR. Although, swMb contains a heme center, it lacks the Fe_B center and nitric oxide reduction activity of NOR. Investigators designed a Fe_B center in swMb by introducing one glutamate and three histidine residues as ligands in the active site of

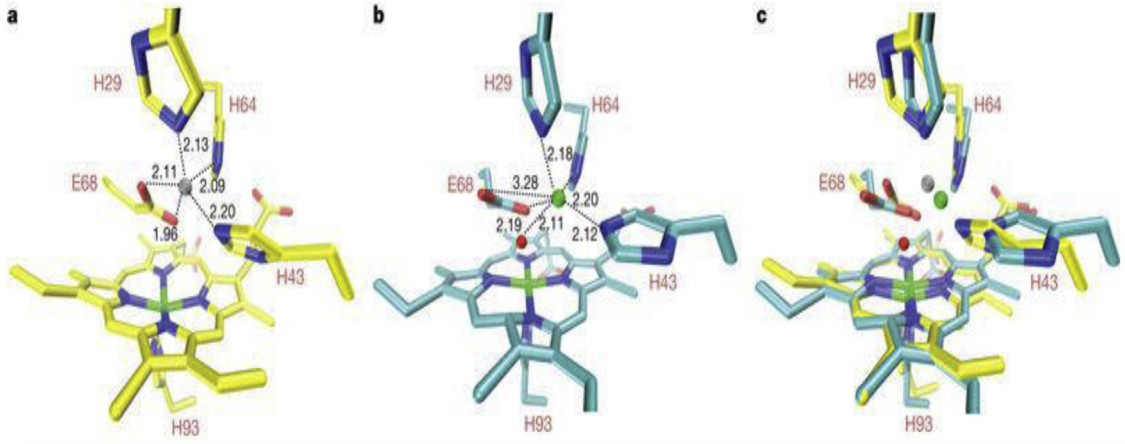

Fig. 6. Superimposition of crystalized rationally designed FeBMb with minimized generated model. (a) Minimized generated model of FeBMb with Zn (II) in the FeB site. (b) Crystal structure of Fe(II)-FeBMb with 1.72 Å resolution. (c) Overlay of FeBMb model (yellow) with Fe(II)-FeBMb crystal structure (cyan). Reprinted with permission [87]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

NOR. Modeling of Fe_BMb in crystal structure of swMb was performed with mutated F43H, L29H and V68E residues employing modeling extension in VMD [88] (Fig. 6). As MD simulation was carried out the energy minimization was performed through NAMD package [87] under 5000 and 10000 minimization steps at 0 and 310 k, respectively for evaluation of appropriate binding of Fe_B site to Fe ion. Further studies of crystalline structure of designed protein revealed glutamate stabilized iron binding to Fe_B site in the active site of NOR. Data analysis demonstrated that the two protons needed for nitric oxide reduction provided by water molecule, histidine and glutamate residues which played significant structural role at the active site as well. It was concluded that glutamate and histidine residues played fundamental roles in stabilizing the iron binding site of Fe_B and NOR activity. In this way a protein model was rationally designed as a similar from the point of structure and function to native NOR. It can be taken as an appropriate model to mechanistic studies of nitric oxide reductase activity [90].

In 2009, Pavlova et al. redesigned haloalkane dehalogenase enzyme of *Rhodococcus rhodochrous* [91]. Interestingly, the mutated enzyme showed 32 fold higher activity toward 1, 2, 3-trichloropropane (TCP) as a substrate in contrast to the native enzyme. The residues which are crucially involved in linking the buried active site to the bulk solvent have been altered by different methods of rational design and directed evolution. Random acceleration molecule dynamic (RAMD) simulations were applied to identify the residues to the altered and simulate the product released through the enzyme's active site [92]. Role of water molecular activity was determined by classical MD. Fig. 7 shows the critical residues identified by saturated mutagenesis (Fig. 7). Improvement of carbon–hologn bond breakage by mutant enzymes was confirmed by kinetic analysis and they obtained up to 32-fold higher activity than the wild type [91].

In an another exciting study, Farasat et al., 2017 [93], increased the affinity of nanobody binding to epidermal growth factor receptor (EGFR) with MD simulation approaches. Five reasonable variants were selected based on measuring desolvation and electrostatic interactions between diverse residues of 7D12 (mutated nanobody) and EGFR. MD simulations for the 50 ns, and pull and umbrella sampling simulation were done and weighted histogram analysis method used for binding free energy calculations. The results showed that A100Q mutation displays an appropriate binding improvement with satisfactory structural stability.

4. Future perspective

As discussed above, MD simulation could be used as a very

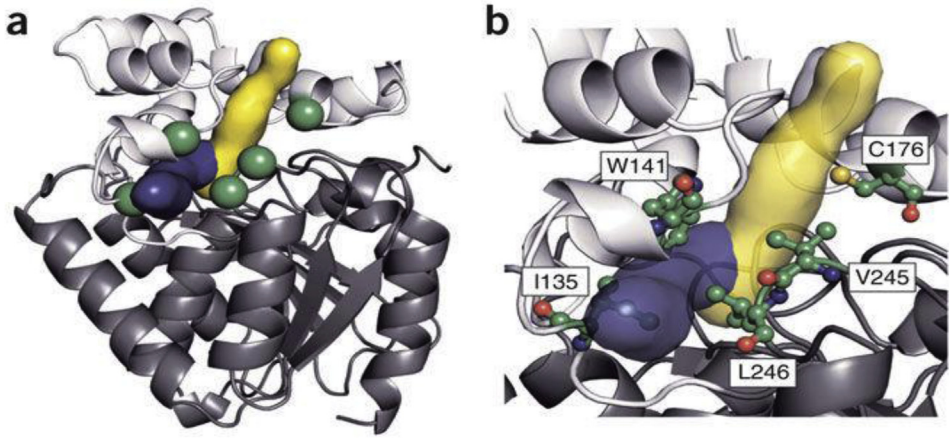

Fig. 7. Selected residues for mutagenesis lining in the hot spot domain of haloalkane dehalogenase enzyme. (a) Cartoon model of wild-type DhaA and (b) Ball and stick model of the residues selected for mutagenesis. Reprinted with permission [88].

beneficial tool for understanding the relation between the protein structure dynamics and its function. Thereby novel designs can be achieved based on the results of such accurate predictions which are inaccessible through the use of static native structures alone. The most important aspect of using MD in protein engineering studies is the ability of such method to provide atomic details of given conformational structure in addition to remarking the changes between native and engineered states that may be unreachable by the other methods. As discussed above by different examples it is obvious that there is a high degree of synergy between MD results and experimental data. So it can be applied to several points in the protein engineering field. With continuing advances in the methodology and the speed of computers, MD studies are being extended to higher systems, greater complex conformational changes and longer time frames. Advanced processor software lead to simulate larger molecules for longer times. Recently, by using high-performance computing systems (HPCs), the simulation time is extended to a range of 100 ns to microseconds, making it possible to study biological phenomena as they happen. With respect to the effect of force field quality on the protein folding; it is expected that the continuous force field improvements, like coarse grained MD simulations, will also lead to more progress.

Since important functions in the cell are often performed by protein complexes which the function is directly dependent on the conformation of the subunits, MD could be a valuable tool for assessing the atomic level of such complex interactions. However, this hadn't been possible unless with dramatic progress have been currently occurred in this field. For example, ATP synthase complex [94,95] and nicotinic acetylcholine receptor [96], have been simulated by molecular dynamics. Interestingly, Grubmüller and coworkers performed MD simulations of complete ribosome structures, which is one of the most important and complex machinery systems in the cell, by combining cry-electron microscopy data and crystallographic structures, at 13 intermediate states of the translation process [97,98]. With continuing advances in computer hardware and algorithms, simulation of other large and complicated systems is expected to be solved in the near future [99]. Surprisingly this method was used to simulate the difficult phenomena such as the formation of phospholipid bilayers in the fusion process of synaptic transmission. Additionally, one of the most valuable aspects of progress in MD simulation method will be their routine use by experimentalists as a tool, for enlightening the interpretation and understanding of data which is now beginning to occur, as illustrated above in different studies. Based on the results available today, it is expectable that the MD will be applied in different fields of biology in order to enhance our understanding of biologic phenomena in the future.

Disclosure of conflict of interest

We have no relevant, potential conflicts of interest to declare.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jmgm.2018.06.009.

References

[1] O. Alvizo, B.D. Allen, S.L. Mayo, Computational protein design promises to revolutionize protein engineering, Biotechniques 42 (1) (2007) 31–35.
[2] R.A. Chica, Protein engineering in the 21st century, Protein Sci. 24 (4) (2015) 431–433.
[3] J. Damborsky, J. Brezovsky, Computational tools for designing and engineering enzymes, Curr. Opin. Chem. Biol. 19 (2014) 8–16.
[4] M.C. Childers, V. Daggett, Insights from molecular dynamics simulations for computational protein design, Molecular Systems Design & Engineering 2 (1) (2017) 9–33.
[5] J. Novotny, 2-Computer Methods in Protein Modeling: an Overview A2-Carey, Paul R, in Protein Engineering and Design, Academic Press, San Diego, 1996, pp. 51–73.
[6] S. Soni, et al., Molecular modeling and molecular dynamics simulations based structural analysis of the SG2NA protein variants, BMC Research Notes 7 (1) (2014) 446.
[7] M.G. Pikkemaat, et al., Molecular dynamics simulations as a tool for improving protein stability, Protein Engineering 15 (3) (2002) 185–192.
[8] M. Karplus, J.A. McCammon, Molecular dynamics simulations of biomolecules, Nat. Struct. Mol. Biol. 9 (9) (2002) 646–652.
[9] W. Wang, et al., Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions, Annu. Rev. Biophys. Biomol. Struct. 30 (1) (2001) 211–243.
[10] T. Hansson, C. Oostenbrink, W. van Gunsteren, Molecular dynamics simulations, Curr. Opin. Struct. Biol. 12 (2) (2002) 190–196.
[11] L. Katz, C. Levinthal, Interactive computer graphics and representation of complex biological structures, Annu. Rev. Biophys. Bioeng. 1 (1) (1972) 465–504.
[12] M. Levitt, A. Warshel, Computer simulation of protein folding, Nature 253 (5494) (1975) 694–698.
[13] A.W. Burgess, H.A. Scheraga, Assessment of some problems associated with prediction of the three-dimensional structure of a protein from its amino-acid sequence, Proc. Natl. Acad. Sci. Unit. States Am. 72 (4) (1975) 1221–1225.
[14] S. Tanaka, H.A. Scheraga, Model of protein folding: inclusion of short-, medium-, and long-range interactions, Proc. Natl. Acad. Sci. Unit. States Am. 72 (10) (1975) 3802–3806.
[15] I. Kuntz, et al., Calculation of protein tertiary structure, JMB (J. Mol. Biol.) 106 (4) (1976) 983–994.
[16] M. Levitt, R. Sharon, Accurate simulation of protein dynamics in solution, Proc. Natl. Acad. Sci. Unit. States Am. 85 (20) (1988) 7557–7561.
[17] A.D. Mackerell, Empirical force fields for biological macromolecules: overview and issues, J. Comput. Chem. 25 (13) (2004) 1584–1604.
[18] M. Karplus, J. Kuriyan, Molecular dynamics and protein function, Proc. Natl. Acad. Sci. U.S.A. 102 (19) (2005) 6679–6685.
[19] F.R. Salsbury, Molecular dynamics simulations of protein dynamics and their relevance to drug discovery, Curr. Opin. Pharmacol. 10 (6) (2010) 738–744.
[20] A.D. MacKerell Jr., et al., All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (18) (1998) 3586–3616.
[21] J.W. Ponder, D.A. Case, Force fields for protein simulations, Adv. Protein Chem. 66 (2003) 27–85.
[22] W.L. Jorgensen, J. Tirado-Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, J. Am. Chem. Soc. 110 (6) (1988) 1657–1666.
[23] C. Oostenbrink, et al., A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6, J. Comput. Chem. 25 (13) (2004) 1656–1676.
[24] E. Krieger, S.B. Nabuurs, G. Vriend, Homology modeling, Meth. Biochem. Anal. 44 (2003) 509–524.
[25] J.M. Harris, N.E. Martin, M. Modi, Pegylation, Clin. Pharmacokinet. 40 (7) (2001) 539–551.
[26] G.M. Watt, et al., Site-specific glycosylation of an aglycosylated human IgG1-Fc antibody protein generates neoglycoproteins with enhanced function, Chem. Biol. (Lond.) 10 (9) (2003) 807–814.
[27] N.V. Katre, The conjugation of proteins with polyethylene glycol and other polymers: altering properties of proteins to enhance their therapeutic potential, Adv. Drug Deliv. Rev. 10 (1) (1993) 91–114.
[28] S. Frokjaer, D.E. Otzen, Protein drug stability: a formulation challenge, Nat. Rev. Drug Discov. 4 (4) (2005) 298.
[29] N. Ceaglio, et al., Novel long-lasting interferon alpha derivatives designed by glycoengineering, Biochimie 90 (3) (2008) 437–449.
[30] A.M. Sinclair, S. Elliott, Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins, J. Pharmaceut. Sci. 94 (8) (2005) 1626–1635.
[31] L. Runkel, et al., Structural and functional differences between glycosylated and non-glycosylated forms of human interferon-β (IFN-β), Pharmaceut. Res. 15 (4) (1998) 641–649.
[32] A. Varki, J.B. Lowe, Biological roles of glycans, Essentials of Glycobiology, second ed., 2009. Chapter 6.
[33] G.W. Hart, et al., Primary structural requirements for the enzymatic formation of the N-glycosidic bond in glycoproteins. Studies with natural and synthetic peptides, J. Biol. Chem. 254 (19) (1979) 9747–9753.
[34] E. Bause, Structural requirements of N-glycosylation of proteins, Studies with proline peptides as conformational probes 209 (2) (1983) 331–336. Biochemical Journal.
[35] M. Samoudi, et al., Rational design of hyper-glycosylated interferon beta analogs: a computational strategy for glycoengineering, J. Mol. Graph. Model. 56 (2015) 31–42.
[36] M. Samoudi, et al., Rational design of glycoengineered interferon-β analogs with improved aggregation state: experimental validation, Protein Eng. Des.

Sel. 30 (1) (2016) 23–30.
[37] S. Elliott, et al., Enhancement of therapeutic protein in vivo activities through glycoengineering, Nat. Biotechnol. 21 (4) (2003) 414.
[38] F. Ghasemi, et al., In silico designing of hyper-glycosylated analogs for the human coagulation factor IX, J. Mol. Graph. Model. 68 (2016) 39–47.
[39] H.S. Lee, Y. Qi, W. Im, Effects of N-glycosylation on protein conformation and dynamics: protein Data Bank analysis and molecular dynamics simulation study, Sci. Rep. (2015) 5.
[40] J. Wang, et al., AMBER 10, University of California, San Francisco, 2008. There is no corresponding record for this reference).
[41] V. Hornak, et al., Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins: Structure, Function, and Bioinformatics 65 (3) (2006) 712–725.
[42] K.N. Kirschner, et al., GLYCAM06: a generalizable biomolecular force field, *Carbohydrates.* Journal of computational chemistry 29 (4) (2008) 622–655.
[43] N.T. Wood, et al., The influence of N-linked glycans on the molecular dynamics of the HIV-1 gp120 V3 loop, PLoS One 8 (11) (2013) e80301.
[44] A. Abuchowski, et al., Effect of covalent attachment of polyethylene glycol on immunogenicity and circulating life of bovine liver catalase, J. Biol. Chem. 252 (11) (1977) 3582–3586.
[45] A. Abuchowski, et al., Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol, J. Biol. Chem. 252 (11) (1977) 3578–3581.
[46] P. Caliceti, F.M. Veronese, Pharmacokinetic and biodistribution properties of poly (ethylene glycol)–protein conjugates, Adv. Drug Deliv. Rev. 55 (10) (2003) 1261–1277.
[47] J.M. Harris, R.B. Chess, Effect of pegylation on pharmaceuticals, Nat. Rev. Drug Discov. 2 (3) (2003) 214.
[48] F.M. Veronese, A. Mero, The impact of PEGylation on biological therapies, BioDrugs 22 (5) (2008) 315–329.
[49] H. Lee, Molecular modeling of PEGylated peptides, dendrimers, and single-walled carbon nanotubes for biomedical applications, Polymers 6 (3) (2014) 776–798.
[50] N. Nischan, C.P. Hackenberger, Site-specific PEGylation of proteins: recent developments, J. Org. Chem. 79 (22) (2014) 10727–10733.
[51] C. Yang, D. Lu, Z. Liu, How PEGylation enhances the stability and potency of insulin: a molecular dynamics simulation, Biochemistry 50 (13) (2011) 2585–2593.
[52] D. Van Der Spoel, et al., GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (16) (2005) 1701–1718.
[53] B. Hess, et al., GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theor. Comput. 4 (3) (2008) 435–447.
[54] W.F. van Gunsteren, et al., Biomolecular Simulation: the {GROMOS96} Manual and User Guide, 1996.
[55] H.J. Berendsen, et al., Interaction Models for Water in Relation to Protein Hydration, in Intermolecular Forces, Springer, 1981, pp. 331–342.
[56] Q. Mu, T. Hu, J. Yu, Molecular insight into the steric shielding effect of PEG on the conjugated staphylokinase: biochemical characterization and molecular dynamics simulation, PLoS One 8 (7) (2013) e68559.
[57] F.M. Veronese, J.M. Harris, Introduction and overview of peptide and protein pegylation, Adv. Drug Deliv. Rev. 54 (4) (2002) 453–456.
[58] S. Elliott, E. Pham, I.C. Macdougall, Erythropoietins: a common mechanism of action, Exp. Hematol. 36 (12) (2008) 1573–1584.
[59] A. Maleki, et al., Design, modeling, and expression of erythropoietin cysteine analogs in Pichia pastoris: improvement of mean residence times and in vivo activities through cysteine-specific PEGylation, Eur. J. Pharm. Biopharm. 80 (3) (2012) 499–507.
[60] J.D. Gans, D. Shalloway, Qmol: a program for molecular visualization on Windows-based PCs, J. Mol. Graph. Model. 19 (6) (2001) 557–559.
[61] R. Fraczkiewicz, W. Braun, Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules, J. Comput. Chem. 19 (3) (1998) 319–333.
[62] R.A. Cohan, et al., Design, modeling, expression, and chemoselective PEGylation of a new nanosize cysteine analog of erythropoietin, Int. J. Nanomed. 6 (2011) 1217.
[63] S. Hadadian, et al., Chemoselective PEGylation of cysteine analogs of human basic fibroblast growth factor (hbFGF)-design and expression, Trop. J. Pharmaceut. Res. 13 (10) (2014) 1601–1607.
[64] H. Mirzaei, et al., Computational and nonglycosylated systems: a simpler approach for development of nanosized PEGylated proteins, Drug Design, Development and Therapy 10 (2016) 1193.
[65] K. Knop, et al., Poly (ethylene glycol) in drug delivery: pros and cons as well as potential alternatives, Angew. Chem. Int. Ed. 49 (36) (2010) 6288–6308.
[66] M. Schlapschy, et al., Fusion of a recombinant antibody fragment with a homo-amino-acid polymer: effects on biophysical properties and prolonged plasma half-life, Protein Eng. Des. Sel. 20 (6) (2007) 273–284.
[67] M. Schlapschy, et al., PASylation: a biological alternative to PEGylation for extending the plasma half-life of pharmaceutically active proteins, Protein Eng. Des. Sel. 26 (8) (2013) 489–501.
[68] R.E. Kontermann, Strategies for extended serum half-life of protein therapeutics, Curr. Opin. Biotechnol. 22 (6) (2011) 868–876.
[69] M.H. Hedayati, et al., Molecular design, expression and evaluation of PASylated human recombinant erythropoietin with enhanced functional properties, Protein J. 36 (1) (2017) 36–48.
[70] F. Khodabakhsh, et al., Development of a novel nano-sized anti-VEGFA nanobody with enhanced physicochemical and pharmacokinetic properties, Artificial Cells, Nanomedicine, and Biotechnology (2017) 1–13.
[71] P.J. Flory, Theory of elastic mechanisms in fibrous proteins, J. Am. Chem. Soc. 78 (20) (1956) 5222–5235.
[72] J. Clarke, A.R. Fersht, Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation, Biochemistry 32 (16) (1993) 4322–4329.
[73] J. Mansfeld, et al., Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond, J. Biol. Chem. 272 (17) (1997) 11152–11156.
[74] J.E. Villafranca, et al., Directed mutagenesis of dihydrofolate reductase, Science 222 (1983) 782–789.
[75] L.J. Perry, R. Wetzel, Disulfide bond engineered into T4 lysozyme: stabilization of the protein toward thermal inactivation, Science 226 (1984) 555–558.
[76] M.W. Pantoliano, et al., Protein engineering of subtilisin BPN': enhanced stabilization through the introduction of two cysteines to form a disulfide bond, Biochemistry 26 (8) (1987) 2077–2082.
[77] J. Thornton, Disulphide bridges in globular proteins, JMB (J. Mol. Biol.) 151 (2) (1981) 261–287.
[78] C. Pabo, E. Suchanek, Computer-aided model-building strategies for protein design, Biochemistry 25 (20) (1986) 5987–5991.
[79] B. Hazes, B.W. Dijkstra, Model building of disulfide bonds in proteins with known three-dimensional structure, Protein Eng. Des. Sel. 2 (2) (1988) 119–125.
[80] R. Sowdhamini, et al., Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis, Protein Eng. Des. Sel. 3 (2) (1989) 95–103.
[81] S.M. Franken, et al., Crystal structure of haloalkane dehalogenase: an enzyme to detoxify halogenated alkanes, EMBO J. 10 (6) (1991) 1297.
[82] J. Zhang, et al., Pentamerization of single-domain antibodies from phage libraries: a novel strategy for the rapid generation of high-avidity antibody reagents, J. Mol. Biol. 335 (1) (2004) 49–56.
[83] K.H.G. Verschueren, et al., Refined x-ray structures of haloalkane dehalogenase at pH 6·2 and pH 8·2 and implications for the reaction mechanism, J. Mol. Biol. 232 (3) (1993) 856–872.
[84] H.-L. Liu, W.-C. Wang, Protein Engineering to Improve the Thermostability of Glucoamylase from Aspergillus Awamori Based on Molecular Dynamics Simulations, vol. 16, 2003, pp. 19–25.
[85] G. Interlandi, et al., Characterization and further stabilization of designed ankyrin repeat proteins by combining molecular dynamics simulations and experiments, J. Mol. Biol. 375 (3) (2008) 837–854.
[86] D. Suplatov, et al., Computational design of a pH stable enzyme: understanding molecular mechanism of penicillin acylase's adaptation to alkaline conditions, PLoS One 9 (6) (2014) e100643.
[87] J.C. Phillips, et al., Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802.
[88] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1) (1996) 33–38.
[89] D. Case, et al., Amber12, University of California, San Francisco, USA, 2012.
[90] N. Yeung, et al., Rational design of a structural and functional nitric oxide reductase, Nature 462 (7276) (2009) 1079–1082.
[91] M. Pavlova, et al., Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate, Nat. Chem. Biol. 5 (10) (2009) 727–733.
[92] P.J. Winn, et al., Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine, Proc. Natl. Acad. Sci. Unit. States Am. 99 (8) (2002) 5361–5366.
[93] A. Farasat, et al., Affinity enhancement of nanobody binding to EGFR: in silico site-directed mutagenesis and molecular dynamics simulation approaches, J. Biomol. Struct. Dyn. 35 (8) (2017) 1710–1728.
[94] J. Ma, et al., A dynamic analysis of the rotation mechanism for conformational change in F1-ATPase, Structure 10 (7) (2002) 921–931.
[95] R.A. Böckmann, H. Grubmüller, Nanoseconds molecular dynamics simulation of primary mechanical energy transfer steps in F1-ATP synthase, Nat. Struct. Biol. 9 (2002) 198.
[96] R.H. Henchman, J.A. McCammon, Structural and dynamic properties of water around acetylcholinesterase, Protein Sci. : A Publication of the Protein Society 11 (9) (2002) 2080–2090.
[97] K.Y. Sanbonmatsu, Computational studies of molecular machines: the ribosome, Curr. Opin. Struct. Biol. 22 (2) (2012) 168–174.
[98] L.V. Bock, et al., Energy barriers and driving forces in tRNA translocation through the ribosome, Nat. Struct. Mol. Biol. 20 (2013) 1390.
[99] J.R. Perilla, et al., Molecular dynamics simulations of large macromolecular complexes, Curr. Opin. Struct. Biol. 31 (2015) 64–74.