Biomacromol. J. Vol. 2, No. 1, 8-20, June 2016.

Biomacromolecular Journal
www.bmmj.org

Protein Stability, Folding, Disaggregation and Etiology of Conformational Malfunctions

A.A. Moosavi-Movahedi[a,b,*], A.A. Saboury[a, b], S. Hosseinkhani[c], A. Lohrasbi-Nejad[d], M. Habibi-Rezaei[e,f], P. Maghami[a], M. Atri[g] and L. Fotouhi[a]

[a]*Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran*
[b]*Center of Excellence in Biothermodynamics, University of Tehran, Tehran, Iran*
[c]*Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran*
[d]*Department of Biotechnology, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran*
[e]*School of Biology, College of Science, University of Tehran, Tehran, Iran*
[f]*Nano-Biomedicine Center of Excellence, Nanoscience and Nanotechnology Research Center, University of Tehran, Tehran, Iran*
[g]*Department of Molecular & Cell Biology, Faculty of Basic Sciences, University of Mazandaran, Babolsar, Iran*
(Received 28 February 2016, Accepted 16 August 2016)

ABSTRACT

Estimation of protein stability is important for many reasons: first providing an understanding of the basic thermodynamics of the process of folding, protein engineering, and protein stability plays important role in biotechnology especially in food and protein drug design. Today, proteins are used in many branches, including industrial processes, pharmaceutical industry, and medical fields. Activity and stability of proteins are essential for providing healthy condition or required during their production, storage and use in their applications. Through the first part of this review, we aim to define the protein stability terms and factors. Any factor induces stabilizing conformation and/or aggregation of proteins might be of importance in etiology of the conformational diseases. In the second part we are going to clarify a comprehensive definition of protein stability issues with special emphasis on the advantages of these concepts in protein conformational diseases and biotechnology with a short insight to protein engineering approaches.

Keywords: Protein stability, Conformational diseases, Reversibility, Refolding, Protein aggregation

DEFINITION OF PROTEIN STABILITY

Protein stability is defined as the net balance of forces which determines whether a protein has its native fold or a denatured state. Before discussing of stability issues in proteins, we must have an exact definition of what we mean by stability. The word is used in different ways by different people. There are two definitions for protein stability; thermodynamic and kinetic stability. Thermodynamic stability is related to denatured state in equilibrium with native protein. Kinetic stability is regarding to significant energy barrier separating native and denatured or partially denatured state. Thermodynamic stability of a protein is defined as the difference in free energy between the native and denatured states under physiological condition [1].

Now, Let us discuss the physical and the chemical stabilities, based on the nature of interactions that are targeted. The physical stability impacts on non-covalent intra-molecular forces as a thermodynamic and equilibrium approach and is defined as the net balance of these forces to determine whether a protein will be in its native folded conformation or a denatured state. It normally refers to the tendency of a protein to preserve its native and functional structure and opposite protein reversible denaturation. Protein denaturation is a process that is caused by various factors, including high and low temperatures, high pressures, ultrasound, high-intensity irradiation (including microwaves), organic solvents, certain salts, detergents and chaotrops. Although, the structure of proteins is enormously diverse, the features that govern protein denaturation are

*Corresponding author. E-mail: moosavi@ut.ac.ir

very similar. When a protein is reversibly denatured, the native (N) and denatured (D) forms of a protein are in equilibrium, which is governed by the denaturation and defined by equilibrium constant, that is determined experimentally under denaturing conditions by measuring catalytic activity or physical characteristics of the protein. Nevertheless, the chemical stability of proteins is defined as the protein resistance to covalent modification in the range of proteolytic cleavage, single or several amino acid residue changes or any other chemical alterations (oxidation, deamidation, reduction, and hydrolysis) which brings about protein irreversible inactivation[2]. Therefore, the chemical instability refers to the extent of the tendency to the formation or breaking of covalent bonds within a protein. More closely, the biological stability is the protein resistance to degradation by proteases in their metabolic pathways in a determined rate in which protein covalent structure is targeted to be changed. This stability defines the *in vivo* half-life of a protein [2].

In the case of irreversible proteins, the rate of unfolding or kinetic stability is considered more important. A protein which is irreversible or kinetically unstable will unfold more rapidly than a kinetically stable protein. In a kinetically unstable protein, the energy barrier to unfolding is relatively low. The parameters affecting stability are the relative free energies of the folded and the transition state on the unfolding pathway [3].

DETERMINATION OF PROTEIN STABILITY

The Pace analysis or linear extrapolation method which first described by Pace [4] is one of the simplest ways for elucidation of protein folding mechanism and measuring the protein conformational stability. This method is mostly used when macromolecule reversible denaturation follows two-state model (N↔D) where N is the native state and D is the denatured state.

However, it is also possible to obtain comparative quantitative estimates of conformational stability for proteins with more complex denaturation mechanism (when intermediate species are present) by this method [4,5]. There are some complications in studying multistate protein folding: the first one is the determination of the intermediates and the second one is the sequence of intermediate emergence. The next complex problem is the exact time of protein folding stages (secondary structure elements) because every single intermediate properties depend on the individual sub-structures [6-9].

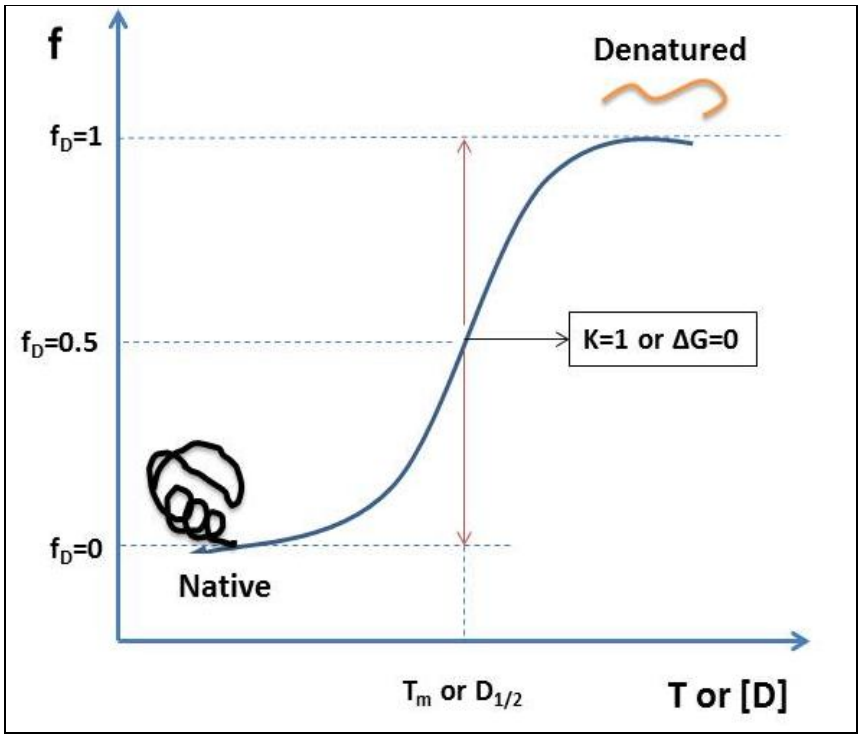
Thermodynamics of Protein Stability

The change in Gibbs free energy, ΔG°, is a measure of conformational stability of native conformation of a globular protein. Comparison of the protein ΔG° under different conditions would be useful for measuring the stability differences accompanied by protein mutation, chemical modification or binding of a specific ligand to the protein. The amount of ΔG^0_{25} (the standard Gibbs free energy of protein denaturation at 25 °C) and $\Delta G^0_{H_2O}$ (the standard Gibbs free energy of protein denaturation in the absence of denaturant) can be quantified using Pace analysis of thermal and chemical denaturation curves, respectively. In addition, we can obtain some information about the mechanism of folding (for example whether unfolding has a two-state or multi-state mechanism) and structure of the protein from the denaturation curve.

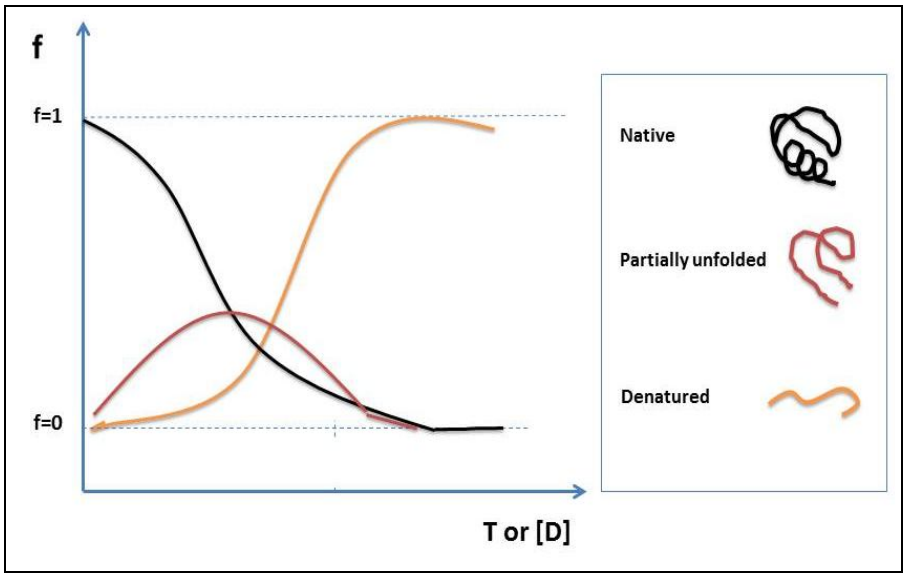
Denaturation Curve Analysis

The extent of denaturation can be detected using perceptible changes of any signal recorded by a conformation-sensitive method such as UV Absorbance, Fluorescence emission, far-UV CD or near-UV CD spectroscopy. Conformational changes are along with chromospheres local environmental changes so they usually result in spectral alterations. A specific wavelength that shows more significant changes is selected to follow unfolding in each method. The signal is measured as a function of temperature or chemical denaturant concentration.

The denaturation curves often show a sigmoidal, cooperative transition from the native to the unfolded (denatured) state. A typical thermal denaturation curve is shown in Scheme 1. It is called ‘‘all-or-none’’ transition. We can divide the denaturation curve to three stages. In the first stage all the protein molecules are native and in the last



Scheme 1. A typical thermal or chemical denaturation curve of a two-state protein



Scheme 2. A typical thermal or chemical denaturation curve of a three-state protein

stage all the molecules are denatured. The second stage is phase transition step in which both native and the denatured states are present in different concentrations. The narration of the scheme 1 for a two-state unfolding can find in the literature [5]. The same is true for chemical denaturation. A similar approach has been suggested for measuring the stability of RNA molecules too [10].

In multi-step denaturation the story is to some extent different. The native state transform to one or more intermediates substantially and at the end a fully denatured state emerges. The view that amino acid strings tend to fold cooperatively *via* two-state structures is mainly based on observations that commonly fail to demonstrate intermediates even when they are present. So if there are methods which are capable of revealing unstable intermediates, the nature of multistate protein can be uncovered [6]. Some recent papers reported the more reliable results for stability estimation of multistate proteins using experimental and calculation techniques [6-9]. In Scheme 2 a three-state denaturation is shown in which one intermediate is produced and transformed to fully unfolded state.

Brief Concern on Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) is commonly used as a technique to investigate the conformational changes of protein upon changing the temperature, which result in gaining thermodynamic parameters, thermal stability and thermal reversibility. Reversibility is a necessary requirement to rigorously measure the thermodynamics of protein stability and thermodynamic analysis and essential for many linked applications like ligand binding analysis and protein concentration dependence among others. The thermodynamic stability of a monomeric folded protein results from the balance of stabilizing and destabilizing forces, which are originated from non-covalent intra-molecular interactions and conformational entropy [11]. Moreover reversibility is the repeatability of the behavior of a protein in a second heating scan by (DSC) [12].

Furthermore, for determining some thermodynamic parameters for protein two-state irreversible system can refer to literature [13].

DENATURATION, INACTIVATION AND AGGREGATION

Protein aggregation is kind of instability that leads to protein conformational diseases such as Alzheimer's, Parkinson, Huntington's and cataract [12,13]. As a result any parameter induced stabilizing native conformation and/or preventing aggregation of proteins might be of importance in etiology of the conformational diseases [13]. Many external factors play critical role in controlling protein aggregation, such as ionic strength, temperature, osmolytes, additives [14-16]. Chemical transformations may result in direct or indirect protein aggregation, such as disulfide bond exchange and non-disulfide cross-linking [17-19].

Although, it is well known that the energy difference between proteins in their native states with unfolded state are very small, but it is often less understood that most proteins are inherently prone to aggregation in their denatured and unfolded or partially unfolded states [20]. Therefore, some conditions *in vivo* or *in vitro* that causes protein denaturation, leading to protein aggregation. The protein aggregates accounts as extremely stable, long-lived, undesirable in biotechnology and harmful in medicine [21]. The undesirable aggregation of unfolded proteins occurs in living biological systems on many occasions, including co-translational misfolding of the synthesized polypeptide chains, mutations, or stress (heat, oxidizing conditions, and toxic compounds) that result in the malfunctioning of cells and organs. The problem of aggregation is challenging not only in medicine cell biology and but also in biotechnological, pharmaceutical, and food industry applications involving proteins. The development of genetic engineering techniques for over expression of proteins has increased attention to the practical aspects of protein folding. Both basic protein science and biotechnology require that the over expressed protein attain the correctly folded conformation.

In the most instances, the reversible denaturation is followed by an irreversible inactivation [22-24]. In irreversible inactivation protein does not recover its native structure and function after removing the denaturing factor or condition; this may be referred as irreversible denaturation or simply inactivation [25].

Protein inactivation is arisen either by chemical or physical mechanisms. In the chemical mechanism, the protein primary structure is changed through: hydrolysis of peptide bonds; oxidation of Cys, Met, His, Tyr and Trp; reduction of S-S bonds, followed by β-elimination; modification of SH groups; deamidation of Asn and Gln residues; racemization of amino acids and finally protein glycation on free amino groups relating to N-terminal residue, Lys residue(s) and Arg residue(s) [26]. In the physical mechanism, the non-covalent interactions are improperly broken or formed. The physical instability (*i.e.* protein unfolding) may facilitate the protein inactivation through chemical or physical mechanisms. Moreover, a chemical event may induce a physical event, for example in case of oxidation which is followed by aggregation. For the instance, proteins are prone to be oxidized through the reaction of certain amino acids with reactive oxygen species (ROS) existing in their environment. Five amino acid residues including methionine (met), cysteine (cys), histidine (his), tryptophan (trp), and tyrosine (tyr) are most susceptible residues in proteins to oxidation [27]. Protein oxidation arises as a consequence of either direct attack by ROS or indirectly through protein glycation or products of lipid peroxidation. In diabetic condition, protein glycation is occurred and many kinds of glycated proteins such as glycated hemoglobin, albumin, and lens crystalline are produced through Schiff base reaction, Amadori rearrangement, and finally advanced glycated end products (AGEs) as well as ROS are generated [28-30]. The generated ROS under diabetic condition causes lipid peroxidation. Lipid peroxidation in turn introduces carbonyl groups into proteins either by direct oxidation of amino acid residues or indirectly by attachment of carbonyl-containing moieties such as 4-hydroxynonenal (4-HNE) or 4-hydroxyhexenal (4-HHE) [31] that may alter physiochemical characteristics and conformation of desired protein, resulting protein aggregation [32]. Direct or indirect oxidation of proteins can increase protein hydrophobicity and enhances protein-protein interactions [33] which follows by aggregation. Nowadays, we know that protein aggregates are threatening the cell viability [33] to bring about aggregation diseases. Aggregated proteins are a significant concern not only in biotechnology for example in biopharmaceutical products, but also in medicine regarding a broad spectrum of protein aggregation diseases known as amyloidosis.

In amyloidosis, insoluble fibrous aggregates of a protein known as amyloid are deposited either systemic or localized as a result of protein instability and sequential changes in protein folding. In systemic amyloidoses deposits may occur in any part of the body, such as AL amyloidosis and TTR-related amyloidosis (ATTR) due to the accumulation of immunoglobulin light chain and transthyretin (TTR) amyloid fibrils, respectively [34]. In local amyloidoses a single organ is affected such as pancreatic aggregation of islet amyloid polypeptide (IAPP) in type 2 diabetes (T2D) and plaque formation in the central nervous system (CNS) due to aggregation of amyloid-β (Aβ) in Alzheimer's disease (AD). Amyloid fibril nomenclature and classification of amyloidosis are based on the identity of the aggregation forming protein [35]. In the XIVth Symposium of the Nomenclature Committee of the International Society of Amyloidosis (ISA) (Symposium of the Society, April 27-May 1, 2014, Indianapolis, IN), 31 extracellular amyloidogenic human proteins were listed. Among the most prevalent amyloidogenic extracellular proteins/peptides, Aβ and IAPP are aggregated in AD and T2D, respectively. Moreover, in spite of a large number of intracellular protein aggregates known as inclusion, few numbers have reported as aggregates with the amyloid features and listed by ISA such as intracellularly tau protein in AD [35].

Various strategies have been reported regarding protein stabilization considering aggregation-derived protein inactivation. Protein aggregation can be prevented using various cosolvents such as chaperons, surfactants, polyols [36] and interfering surfaces [37] through protein stabilization via enhancing intramolecular interactions [38, 39] or destabilizing protein-protein interactions [40] to suppress the intermolecular interactions that lead to protein aggregation [41]. Moreover, the aggregation suppression has been achieved through a hydrophobic competition strategy which has been resulted in the adsorptive interactions between accessible hydrophobic surfaces of the protein refolding intermediates and extended hydrophobic surfaces of the substituted alkyl chains to result active immobilized protein [37].

REVERSIBILITY AND REFOLDING

The process of refolding involves the conversion of non-functional protein aggregates to their functional native state. Molecular chaperones and co-chaperones perform diverse cellular functions. They are involved in the folding and refolding of nascent proteins, inhibition of protein aggregation, refolding of denatured proteins, and assisting the targeting of proteins for degradation by the proteasome and lysosomes. Chaperones can protect the nonnative proteins by binding hydrophobic unraveled or misfolded surfaces, thereby preventing them from interaction with each other or with other proteins in nonproductive or damaging ways. Moreover, protein aggregates can be solubilized and their non-native conformers are able to correctly refold by chaperones [42-46]. A practical and relatively simple approach (compared with using chaperone systems) to solving the aggregation problem is the utilization of low molecular weight "artificial chaperones". This approach was inspired by the mechanisms of action of natural chaperones which often find in vitro to prevent aggregation and increase efficient refolding of denatured proteins in presence of chemical compounds, including denaturants (typically guanidine, urea, and detergents), cyclodextrins, poly(ethylene glycol) [47-49], and surfactants [50]. The efficiency of chaperone-mediated processes are affected by protein concentration; the nature of the protein substrate; size, solubility, and aggregate types; environmental conditions, such as ionic strength, pH, salt type concentration, and temperature; co-solutes; and preservatives. Moreover, because proteins are diverse in structure-functional peculiarities, additives that work well for a particular protein may not function universally [51, 52]. In search for better additives naturally occurring non-denaturing reagents, such as osmolytes [53,54], diamines [55], polyamines [56], amino acids, and their derivatives [57,58] have been used. Arginine is a low molecular weight compound widely used as one of the most effective additive in promoting refolding of aggregated proteins, suppressing aggregation, and enhancing the solubility of aggregatin-prone unfolded molecules [59-63]. It has been reported the effect of a few osmolytes/chemical chaperones, such as glycerol, dimethylsulfoxide, dimethylsulfoxide, trimethyl-amine-N-oxide (TMAO) and ethylene glycol (EG), trehalose, and proline, on the refolding of a Escherichia coli protein maltodextrin glucosidase (MalZ), a 69 kDa monomeric protein responsible for the conversion of maltodextrins to maltose by eliminating one glucose residue from the reducing end at each time [64]. It has been reported that surfactants work as promoters for the refolding of protein (*e.g.,* carbonic anhydrase II) through the formation of soluble folding intermediates and not by the dissolution of aggregates [65]. The SDS exhibits different roles at low concentrations upon interaction with proteins. SDS probably induces the folding or partial unfolding for proteins due to protein characterization. That was reported the dual behavior of SDS as suppressor or enhancer aggregation of insulin and chaperone-like activity of camel _S1-casein [66]. It was also reported that at low concentrations of n-alkyl sulfates, molten globule-like state of cytochrome c is formed. It is expressed the molten globule state of cytochrome c induced by n-alkyl sulfates at low concentrations is a stabilized form with higher free energy in the absence of ligand, $\Delta G(H_2O)$ and m-values. The greater free energy and m-values correspond to the hydrophobic chain length of n-alkyl sulfates, which play a salt-like role exposing the molten globule state. The hydrophobic salts show a greater affinity for the molten globule state, nearly 50 times of the salts without hydrophobic chains. In this work provides the evidence for the stabilization of the molten globule state of cytochrome c by n-alkyl sulfates as salts with hydrophobic chains. Accordingly, the authors label n-alkyl sulfates as hydrophobic salts [67]. Aspergillus niger catalase could be activated up to 180% by a low concentration of SDS at pH 6.4 because of the compaction (folding) of the catalase-SDS complexes under these circumstances. The thermodynamic and hydrodynamic results show the folding of the catalase during the optimum activation at 2 mM SDS [68]. Furthermore SDS at low concentrations ($\leq$ 1 mM) induces the compaction of hemoglobin [43] and histone H1 is also folded by SDS at low concentrations (< 0.5 mM), which was confirmed using various techniques including binding enthalpy curve, calorimetry, polyacrylamide gel electrophoresis, protein titration, and viscometry [69-71].

One of the more effective class of compounds used as potential anti-aggregation agents are the cyclodextrins, toroidal oligosaccharides made up of glucose units. The

hydrophobic interiors of such molecules enables them to sequester aromatic or, to some extent, aliphatic residues on the protein surface that have important role in protein aggregation [72,73]. As it was reported the role of β -cyclodextrin as an anti-aggregation agent on lysozyme-SDS system which seems to prohibit precipitation by changing the nature and the mole fraction (mostly broadening the distribution) of intermediate states [74]. In contrast to single-domain proteins, unfolding of larger multi-domain proteins is often irreversible. Nevertheless, it has been reported that there are significant differences between the proteins with respect to their ability to reversible unfolding transitions [75]. Moreover, there is various strategies have been developed, first, to solubilize the inclusion bodies and in subsequent steps to transfer the proteins into conditions which provide high folding yields. In particular the latter one was studied in more detail by applying various systems including artificial chaperone systems (detergent, cyclodextrins) co-solvent assisted folding (sugars, osmolytes), and in vitro chaperonin systems (GroEL/GroES) [76]. Most of the above mentioned approaches are also considered effective for the reversibility of induced unfolding transitions.

Osmolytes are co-solvents that are used to protect organisms from denaturing condition of harsh environmental stresses. These molecules do not interact with the protein directly, but by changing the surrounding water molecule arrangment and hence the protein-solvent interactions [77]. Their effect seems to be general for all proteins. Osmolytes have no inhibitory or enhancing effects on biological activity of protein under physiological conditions hence are called compatible osmolyte [78,79]. Stabilizing osmolytes include diverse chemical classes as polyols (xylitol, glycerol, sorbitol, adonitol and mannitol), methylamine compounds, certain amino acids and their derivatives, and carbohydrates. Among these chemical classes polyhydric alcohols (polyols) are the most favorable molecules used by nature to protect organisms against the denaturing conditions of high osmotic pressure and freezing [80,81]. They have also been found to be effective stabilizers of proteins and biological assemblies when added at high concentrations [54,82-85].Moreover polyol osmolytes can be used to stabilize proteins at all pH values [86]. In addition osmolytes, such as the small molecules betaine, proline, trehalose, glycerol, dimethylsulfoxide, trimethylamine-N-oxide (TMAO) and ethylene glycol (EG), have been reported to protect native proteins from heat denaturation and favor the formation of native protein oligomers [80,87-89]. Some osmolytes behave as chemical chaperones by promoting the correct folding of unfolded protein *in vitro* and in vivo [82-84,90] ; for example, proline behaves as a protein folding chaperone [91]. Effect of various classes of compatible as well as noncompatible osmolytes on the enzymatic activity, disaggregation, and thermal stability of bovine liver catalase has been investigated. Compatible osmolytes, xylitol, proline, and valine destabilize the unfolded form of the enzyme and, so they increase its disaggregation and thermal stability. This increase in the (thermal) stability is accompanied with a little increase of activity when comparing to the native enzyme at 25 °C. On the other hand, histidine, a non-compatible osmolyte stabilizes the denatured form of the protein and hence causes an overall decrease in the thermal stability and enzymatic activity of the enzyme. Among carbohydrates, trehalose was reported to form a metastable state and functional form of recombinant interferon beta-1b (INTF beta-1b). This study implies that during the refolding process of INTF beta-1b, intermediate structure of the protein is formed. Stability of this protein was improved efficiently by the addition of trehalose as an osmolyte. The hydration shell of INTF beta-1b absorbs more water molecules in the presence of trehalose and a new meta-stable conformation of protein is formed. The trehalose-conserved form of INTF beta-1b could facilitate the post-refolding processes of the protein [92].

GENETIC APPROACHES IN PROTEIN ENGINEERING

According to widespread applications of proteins in many areas of biotechnology and susceptibility of them to instability when transferring proteins from their natural environment into a different one, researchers have looked for methods to increase enzyme performance in new surroundings.

Given that the widespread use of enzymes in a significant number of industries such as food processing, drug discovery, medical application and laundry detergent,

producing them in large scale is of the utmost importance. Today, with the assistance of protein engineering methods recombinant proteins can be designed for specific tasks. The optimization and adaptation of enzymes into environmental conditions are required for commercial purposes. Some of the limitations of enzymes as biocatalysts are their low productivity, low stability, and narrow range of substrates [93,94].

Rational Design

While the protein structure is known and enough knowledge around protein function is existed rational approaches based on amino acid substitution using site-directed mutagenesis have can be used for protein modification.

Development in protein engineering fields has improved in vitro protein functions. For instance, glucose isomerase from *Thermotoga neapolitana* requires neutral pH and high temperature (95 °C) for maximum activity. However, industrial production of corn syrup with a high concentration of fructose requires glucose isomerization to happen at 60 °C, but this enzyme can utilize only 10% of its maximum activity [95]. Glucoamylase is another enzyme that is widely used in the food industry including starch processing to sugar [96]. Given the importance of this enzyme, produced a mutant that had one disulfide bond more than the native enzyme. Firefly luciferase is considered as an unstable protein and often loses its activity at room temperature because of structural changes according to structure-function relationship. Therefore, their enzymatic application is limited due to the structural and functional stability of this protein. Different strategies have been employed to increase the stability of this enzyme such as addition of covalent cross-links (disulfide bonds). It was found that not only the single disulfide bond containing mutant was structurally, but also functionally (the enzymatic activity was 7 fold greater than wild type) more stable than the mutant protein containing two disulfide bonds. The A296C-A326C mutation also increased the reversibility and disaggregation of the protein. Bioinformatics applications showed that the enhanced activity of the single disulfide bond mutant protein was contributed to the expansion of its active site cleft [97].

Moreover, other approaches including reduction in surface loop flexibility [105-107], surface Arginine saturation [108], increase of c-domain rigidity [109], deletion of unusual residues based on distribution in un allowed region of Ramachandran plot [110] and luciferase protection against protease digestion [111-112] have been successfully applied for increase of firefly luciferase thermostability.

Directed Evolution

One method for changing protein structure is directed evolution. Directed evaluation is a general term for all strategies that can create genetic modifications and produce a protein with desirable function using random procedures [98]. Since the 1970s, importance and potential of this approach have been investigated extensively [99-103]. Protein evaluation includes designing and generating a library of desired gene sequences, screening, and selection of mutant proteins that have higher efficiency [104]. Desired gene can be mutated in three ways: (i) random point mutagenesis with the assistance of chemical mutagenesis or error-prone PCR [105,106] (ii) insertion and deletion [107] (iii) gene recombination by DNA shuffling [107].

DNA Shuffling

In recent years, the use of DNA shuffling has been expended. In this approach, a group of homologous genes that have a relatively similar sequence or a single gene that has several mutated sequences is cleaved into small fragments using DNase I. These fragments are then purified and reassembled in a primer-less PCR reaction. In principle, randomly overlapping fragments act as primers and DNA polymerase extends these strands using free 3'-OH terminal. Afterward, a PCR reaction is performed in the presence of primers to generate full-length chimeric genes. PCR amplifications produce a diverse library consisted of recombinant genes that are suitable to be cloned into the expression vector. At last, by screening the clones, the best protein product of these mutant genes is selected [108]. The success of this method depends on size, quality, and diversity of the used library.

According to the development of this technique, the researchers made minor changes to modify the original protocol. For example, Kikuchi and co-workers substituted a mixture of restriction endonuclease instead of DNase I

[109]. Use of this enzymatic mixture led to the production of high-efficiency chimeric genes, so that, DNA strands contain fragments belonging to more than one parent gene. Zhao and co-workers expanded Another type of DNA shuffling [110]. In this case, fragmentation of the parent gene is not required. This method was named staggered extension process (StEP).

DNA shuffling is used as a powerful approach to improving protein folding and protein function [111,112]. Green fluorescent protein (GFP) is a unique protein with the fluorescence light-emitting characteristics widely used in biological research. The first effort to improve the function of GFP as a fluorescence probe was made by Pang and co-workers [113]. They showed that changing the codon at position 65 leads to create a mutant (S65T) that shows enhanced brightness, faster chromophore formation and slower photobleaching. DNA shuffling method have been used to improve the performance of GFP [111]. They could obtain a mutant with a performance 18 times more than that of the wild-type GFP.

DNA shuffling, as a strong method, has been very effective in the field of enzyme stabilization. It was reported the half-life of subtilisin S41 increased 1200 times more than that of the wild-type protein at 60 °*C*. Furthermore, they showed that melting temperature of the mutant has increased by about 25 °*C* compared to the native protein [114-116]. L-Arabinose isomerase is an enzyme which has recently attracted commercial interest due to isomerization between D-galactose and D-tagatose [117]. According to the use of tagatose as a sugar substitute, many attempts have been made to improve the performance of L-Arabinose isomerase [118]. Producing tagatose from lactose is performed in two steps using different enzymes. In the first step, lactose hydrolyzes by β-galactosidase activity at pH 6-7. In the second step, L-Arabinose isomerase converts galactose to tagatose at optimum pH of 8-8.5. To creating a single-step reaction, an enzyme with a lower optimum pH is required [119]. Change of pH-activity profile of arabinose isomerase was reported through directed evolution [120], so that; the optimal pH of created mutant (Q408V and R408V) was shifted to pH 7.5.

Moreover, DNA shuffling method was used to enhance catalytic properties of phytase from *Aspergillus niger* 113 [121]. They measured kinetic parameters belonged to a mutant enzyme (K41E, E121F) and explained that specific activity of these mutant has increased 2.5 and 3.1 times as compared to the wild-type protein.

In comparison with site-directed mutagenesis, DNA shuffling method is used when the molecular basis of the enzyme is poorly understood. The large library consisted of chimeric and mutated hybrid genes is generated by DNA shuffling approach and screened for the most improved mutated protein [122,123].

Error-prone PCR

The ability of directed evaluation in creating a synthetic molecule with new physical and functional properties depends on creating the high-quality library. The combinatorial genetic library can be designed and produced by error-prone PCR method [124]. To increase the mutation rate in a single gene, PCR is performed under specific conditions that reduce the fidelity of DNA polymerase. These special circumstances include: (1) increased the concentration of Taq DNA polymerase, (2) increased concentration of Mg^{2+}, (3) increased polymerase extension time, (4) increased concentration of dNTP and (5) using Mn^{2+} in the reaction [125,126].

Another variation of this technique refers to the procedures whereby a wild-type sequence cloned into the vector, at the first step. Error-prone PCR then amplifies the whole plasmid. In this approach, after mutagenesis of the inserted gene the next steps such as digestion of DNA with restriction enzymes, extraction of DNA from agarose gel, and ligation of the gene into the vector are not required [127]. This method is named error-prone rolling circle amplification (error-prone RCA) and results in a randomly mutated vector library with 3-4 mutations per kilobase [127].

Lipase is an enzyme widely used in commercial applications. So far, several studies have been done on stabilization of the enzyme. stabilization of *bacillus* lipase via ep-PCR has been reported [128]. When Ile replaced with Thr, the activity of mutant lipase was improved compared with the wild-type enzyme. Thus, the half-life of the mutant lipase was enhanced from 7 min to 21 min at 50 °C. Furthermore, they have reported that this substitution in the

protein structure has increased the k_{cat} and k_{cat}/K_m of mutant 2 and 5 times as compared with the wild-type lipase, respectively [128]. It was reported improving of stability of lipase belonging to *Rhizopus niveus*. To achieve this goal, they used ep-PCR and DNA-shuffling techniques. Their results demonstrated optimum temperature for activity of produced mutant using ep-PCR, which had three mutations in its amino acid sequence (P18H, A36T, and E218V), was 15 °C higher than the optimum temperature for native enzyme. After that, they constructed chimeric lipase 1 (CL-1; P18H and A36T) and chimeric lipase 2 (CL-2; E218V) with the assistance of DNA shuffling between mutant and wild-type genes. They observed that optimum temperature for CL-1 activity was similar to that of the wild-type, whereas optimum temperature for CL-2 activity was higher compared to mutated lipase [145].

CONCLUSIONS

In conclusion, protein malfunction contributes to pathogenesis of a wide range of diseases, and this condition happen as conformational or physical instability which is triggered by various destabilizing conditions such as oxidation and glycation. On the other hand it has attained great importance as a common problem encountered during manufacture and storage of proteins, especially in biopharmaceuticals and food science since they are associated with decreased bioactivity and increased immunogenicity. Therefore, adoption of a preventive strategy to confront with protein aggregation is growing concern both in bio-industry and medicine. Thus, understanding the term of protein stability and protein folding may contribute to have brighter insight into mechanisms involved and development of ways to prevent protein inactivation and aggregation and also contribute to the treatment of these diseases or create better strategies to produce more useful enzymes. On the other hand, the optimization and adaptation of enzymes to variable conditions can be done with the assistance of protein engineering methods such as DNA shuffling and error-prone PCR. With the help of recombinant proteins we may overcome the limitation of enzymes applicability as the biocatalysts.

ACKNOWLEDGEMENTS

The support of University of Tehran, Center of Excellence in Biothermodynamics (CEBiotherm), Iran National Science Foundation (INSF), Iran National Elites Foundation (INEF) and UNESCO Chair in Interdisciplinary Research in Diabetes, Iran Society of Biophysical Chemistry are gratefully acknowledged.

REFERENCES

[1] J.M. Sanchez-Ruiz, Biophys. Chem. 148 (2010) 1.
[2] M.C. Manning, K. Patel, R.T. Borchardt, Pharm. Res. 6 (1989) 903.
[3] K.P. Murphy, Protein Structure, Stability, and Folding (2001) 1.
[4] C. Pace, Methods Enzymol. 131 (1986) 266.
[5] C.N. Pace, Trends Biotechnol. 8 (1990) 93.
[6] L. Mayne, S.W. Englander, Protein Sci. 9 (2000) 1873.
[7] T.N. Melnik, T.V. Povarnitsyna, A.S. Glukhov, B.S. Melnik, PloS one 7 (2012) e48604.
[8] J.T. Huang, J.P. Cheng, Proteins: Structure, Function, and Bioinformatics 72 (2008) 44.
[9] J. Schönfelder, R. Perez-Jimenez, V. Muñoz, Nature Communications 7 (2016).
[10] V.M. Shelton, T.R. Sosnick, T. Pan, Biochemistry 38 (1999) 16831.
[11] S.K. Haq, R.H. Khan, Int. J. Biol. Macromol. 35 (2005) 111.
[12] E.H. Koo, P.T. Lansbury, J.W. Kelly, Proceedings of the National Academy of Sciences 96 (1999) 9989.
[13] W. Wang, Int. J. Pharm. 289 (2005) 1.
[14] A.P. Ben-Zvi, P. Goloubinoff, J. Struct. Biol. 135 (2001) 84.
[15] M. Kosmaoglou, N. Schwarz, J.S. Bett, M.E. Cheetham, Prog. Retin. Eye Res. 27 (2008) 434.
[16] T. Sarkar, G. Mitra, S. Gupta, T. Manna, A. Poddar, D. Panda, K.P. Das, B. Bhattacharyya, Eur. J. Biochem. 271 (2004) 1488.
[17] D. Whitley, S.P. Goldberg, W.D. Jordan, J. Vasc. Surg.29 (1999) 748.
[18] H. Grallert, J. Buchner, J. Struct. Biol. 135 (2001) 95.
[19] P. Rodriguez, J. Pelletier, G.B. Price, M. Zannis-

Hadjopoulos, J. Mol. Biol. 298 (2000) 225.
[20] C.J. Roberts, Trends Biotechnol. 32 (2014) 372.
[21] M. Stefani, Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease 1739 (2004) 5.
[22] R.K.O. Apenten, N. Berthalon, Food Chem. 51 (1994) 15.
[23] A. Jakób, J. Bryjak, H. Wójtowicz, V. Illeová, J. Annus, M. Polakovič, Food Chem. 123 (2010) 369.
[24] R. Lumry, H. Eyring, J. Phys. Chem. 58 (1954) 110.
[25] M. Zlokarnik, M. Zlokarnik, Ullmann's Encyclopedia of Industrial Chemistry Wiley-VCH Verlag GmbH & Co. KGaA, 2008.
[26] P. Rondeau, E. Bourdon, Biochimie 93 (2011) 645.
[27] S. Li, C. Schöneich, R.T. Borchardt, Biotechnol. Bioeng. 48 (1995) 490.
[28] M. Goodarzi, A. Moosavi-Movahedi, M. Habibi-Rezaei, M. Shourian, H. Ghourchian, F. Ahmad, M. Farhadi, A. Saboury, N. Sheibani, Spectrochimica Acta Part A: Mol. Biomol. Spectroscopy 130 (2014) 561.
[29] M. Habibi Rezaei, M. Bakhti, A.A. Moosavi-Movahedi, S.J. Zargar, H. Ghorchian, Biomacromol. J. 1 (2015) 212.
[30] M.R. Khazaei, M. Habibi-Rezaei, F. Karimzadeh, A.A. Moosavi-Movahedi, A.A. Sarrafnejhad, F. Sabouni, M. Bakhti, J. Biochem. 144 (2008) 197.
[31] G. Barrera, S. Pizzimenti, E.S. Ciamporcero, M. Daga, C. Ullio, A. Arcaro, G.P. Cetrangolo, C. Ferretti, C. Dianzani, A. Lepore, Antioxid. Redox Sign. 22 (2015) 1681.
[32] H. Mirzaei, F. Regnier, J. Chromatogr. B 873 (2008) 8.
[33] R.S. Sohal, Free Radical Biol. Med. 33 (2002) 37.
[34] A. Aguzzi, T. O'Connor, Nat. Rev. Drug Discov. 9 (2010) 237.
[35] J.D. Sipe, M.D. Benson, J.N. Buxbaum, S.-I. Ikeda, G. Merlini, M.J. Saraiva, P. Westermark, Amyloid 21 (2014) 221.
[36] W. Wang, S. Nema, D. Teagarden, Int. J. Pharm. 390 (2010) 89.
[37] F. Nabati, M. Habibi-Rezaei, M. Amanlou, A. Moosavi-Movahedi, J. Mol. Catal. B-Enzym. 70 (2011) 17.
[38] A. Neagu, M. Neagu, A. Dér, Biophys. J. 81 (2001) 1285.
[39] F. Edwin, Y.V. Sharma, M. Jagannadham, Biochem. Biophys. Res. Commun. 290 (2002) 1441.
[40] V. Vagenende, M.G. Yap, B.L. Trout, Biochemistry 48 (2009) 11084.
[41] S.E. Bondos, A. Bicknell, Anal. Biochem. 316 (2003) 223.
[42] M. Beissinger, J. Buchner, Biol. Chem. 379 (1998) 245.
[43] J. Frydman, Annu. Rev. Biochem. 70 (2001) 603.
[44] H. Wegele, L. Müller, J. Buchner, Hsp70 and Hsp90-a relay team for protein folding, Reviews of physiology, biochemistry and pharmacology, Springer2004, pp. 1.
[45] A. Mogk, B. Bukau, Curr. Biol. 14 (2004) R78.
[46] C.-P. Liu, S. Perrett, J.-M. Zhou, J. Biol. Chem. 280 (2005) 13315.
[47] D. Rozema, S.H. Gellman, J. Am. Chem. Soc. 117 (1995) 2373.
[48] X.Y. Dong, J.H. Shi, Y. Sun, Biotechnol. Prog. 18 (2002) 663.
[49] F. Wang, Y. Liu, J. Li, G. Ma, Z. Su, J. Chromatogr. 1115 (2006) 72.
[50] D. Lu, Z. Liu, J. Wu, J. Chem. Phys. 126 (2007) 064906.
[51] S. Diamant, A.P. Ben-Zvi, B. Bukau, P. Goloubinoff, J. Biol. Chem. 275 (2000) 21107.
[52] A. Puig, H.F. Gilbert, J. Biol. Chem. 269 (1994) 7764.
[53] A. De Marco, L. Vigh, S. Diamant, P. Goloubinoff, Cell Stress Chaperon. 10 (2005) 329.
[54] Y.S. Kim, L.S. Jones, A. Dong, B.S. Kendrick, B.S. Chang, M.C. Manning, T.W. Randolph, J.F. Carpenter, Protein Sci. 12 (2003) 1252.
[55] M. Okanojo, K. Shiraki, M. Kudou, S. Nishikori, M. Takagi, J. Biosci. Bioeng. 100 (2005) 556.
[56] M. Kudou, K. Shiraki, S. Fujiwara, T. Imanaka, M. Takagi, Eur. J. Biochem. 270 (2003) 4547.
[57] S. Taneja, F. Ahmad, Biochem. J. 303 (1994) 147.
[58] K. Shiraki, M. Kudou, R. Sakamoto, I. Yanagihara, M. Takagi, Biotechnol. Prog. 21 (2005) 640.
[59] T. Arakawa, K. Tsumoto, Biochem. Biophys. Res. Commun. 304 (2003) 148.
[60] K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, T.

Arakawa, Biochem. Biophys. Res. Commun. 312 (2003) 1383.
[61] K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J.S. Philo, T. Arakawa, Biotechnol. Prog. 20 (2004) 1301.
[62] K. Shiraki, M. Kudou, S. Nishikori, H. Kitagawa, T. Imanaka, M. Takagi, Eur. J. Biochem. 271 (2004) 3242.
[63] B.M. Baynes, D.I. Wang, B.L. Trout, Biochemistry 44 (2005) 4919.
[64] S. Paul, S. Punam, T.K. Chaudhuri, FEBS J. 274 (2007) 6000.
[65] D. Wetlaufer, Y. Xie, Protein Sci. 4 (1995) 1535.
[66] J. Badraghi, A.A. Moosavi-Movahedi, A.A. Saboury, R. Yousefi, A. Sharifzadeh, J. Hong, T. Haertlé, A. Niasari-Naslaji, N. Sheibani, Int. J. Biol. Macromol. 45 (2009) 511.
[67] A. Moosavi-Movahedi, J. Chamani, Y. Goto, G. Hakimelahi, J. Biochem. 133 (2003) 93.
[68] A. Moosavi-Movahedi, Pure Appl. Chem. 66 (1994) 71.
[69] S. Bathaie, A. Moosavi-Movahedi, B. Ranjbar, A. Saboury, Colloids Surf. B. Biointerfaces 28 (2003) 17.
[70] A. Moosavi-Movahedi, J. Iran. Chem. Soc. 2 (2005) 189.
[71] A. Moosavi-Movahedi, A. Rabbani, M. Godarzi, B. Goliaei, Thermochimica Acta 154 (1989) 205.
[72] P. Camilleri, N.J. Haskins, D.R. Hewlett, FEBS Lett. 341 (1994) 256.
[73] D.E. Otzen, B.R. Knudsen, F. Aachmann, K.L. Larsen, R. Wimmer, Protein Sci. 11 (2002) 1779.
[74] P. Pirzadeh, A. Moosavi-Movahedi, B. Hemmateenejad, F. Ahmad, M. Shamsipur, A. Saboury, Colloids Surf. B. Biointerfaces 52 (2006) 31.
[75] K. Strucksberg, T. Rosenkranz, J. Fitter, Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1774 (2007) 1591.
[76] C. Lange, R. Rudolph, Protein Folding Handbook (2005) 1245.
[77] P.H. Yancey, M.E. Clark, S.C. Hand, R.D. Bowlus, G.N. Somero, Science 217 (1982) 1214.
[78] S.N. Timasheff, Annu. Rev. Biophys. Biomol. Struct. 22 (1993) 67.
[79] P.H. Yancey, Sci. Prog. 87 (2004) 1.
[80] J. Carpenter, S. Prestrelski, T. Arakawa, Arch. Biochem. Biophys. 303 (1993) 456.
[81] S. Prestrelski, T. Arakawa, J. Carpenter, Arch. Biochem. Biophy. 303 (1993) 465.
[82] S.Y. Gerlsma, E.R. Stuur, Int. J. Pept. Protein research 4 (1972) 377.
[83] G. Xie, S.N. Timasheff, Protein Sci. 6 (1997) 211.
[84] G. Xie, S.N. Timasheff, Protein Sci. 6 (1997) 222.
[85] J.K. Kaushik, R. Bhat, J. Phys. Chem. B 102 (1998) 7058.
[86] I. Haque, R. Singh, A.A. Moosavi-Movahedi, F. Ahmad, Biophys. Chem. 117 (2005) 1.
[87] L.J. Borowitzka, A.D. Brown, Arch. Microbiol. 96 (1974) 37.
[88] A. Brown, J.R. Simpson, Microbiology 72 (1972) 589.
[89] M.B. Burg, P.F. Kador, J. Clin. Invest. 81 (1988) 635.
[90] S.D. Edmands, K.S. Hughs, S.Y. Lee, S.D. Meyer, E. Saari, P.H. Yancey, Kidney Int. 48 (1995) 344.
[91] D. Samuel, T.K.S. Kumar, G. Ganesh, G. Jayaraman, P.-W. Yang, M.-M. Chang, V.D. Trivedi, S.-L. Wang, K.-C. Hwang, D.-K. Chang, Protein Sci. 9 (2000) 344.
[92] A. Fazeli, S.A. Shojaosadati, M.R. Fazeli, K. Khalifeh, S. Ariaeenejad, A.A. Moosavi-Movahedi, J. Biotechnol. 163 (2013) 318.
[93] J.D. Rozzell, Bioorgan. Med. Chem. 7 (1999) 2253.
[94] H.E. Schoemaker, D. Mink, M.G. Wubbolts, Science 299 (2003) 1694.
[95] S.B. Rubin-Pitel, H. Zhao, Comb. Chem. High T. Scr. 9 (2006) 247.
[96] P. Kumar, T. Satyanarayana, Crit. Rev. Biotechnol. 29 (2009) 225.
[97] M. Naderi, A.A Moosavi-Movahedi, S. Hosseinkhani, M. Nazari, M. Bohlooli, J. Hong, H. Hadi-Alijanvand, N. Sheibani, Protein Peptide Lett. 22 (2015) 23.
[98] L. Yuan, I. Kurek, J. English, R. Keenan, Microbiol. Mol. Biol. Rev. 69 (2005) 373.
[99] E.T. Farinas, T. Bulter, F.H. Arnold, Curr. Opin. Biotechnol. 12 (2001) 545.
[100] M.T. Reetz, M. Rentzsch, A. Pletsch, M. Maywald,

CHIMIA Int. J. Chem. 56 (2002) 721.
[101] S.V. Taylor, P. Kast, D. Hilvert, Angew. Chem. Int. Edit. 40 (2001) 3310.
[102] G.S. Waldo, Curr. Opin. Chem. Biol. 7 (2003) 33.
[103] H. Tao, V.W. Cornish, Curr. Opin. Chem. Biol. 6 (2002) 858.
[104] M.F. Cole, E.A. Gaucher, Curr. Opin. Chem. Biol. 15 (2011) 399.
[105] O. Kuchner, F.H. Arnold, Trends Biotechnol. 15 (1997) 523.
[106] S. Sen, V.V. Dasu, B. Mandal, Appl. Biochem. Biotechnol. 143 (2007) 212.
[107] D.D. Jones, Nucleic Acids Res. 33 (2005) e80.
[108] S. Lutz, U.T. Bornscheuer (2009).
[109] M. Kikuchi, K. Ohnishi, S. Harayama, Gene 236 (1999) 159.
[110] H. Zhao, L. Giver, Z. Shao, J.A. Affholter, F.H. Arnold, Nat. Biotechnol. 16 (1998) 258.
[111] A. Crameri, E. Whitehorn, E. Tate, W. Stemmer, A. Crameri, P. Kitts, P. Kitts, Nat. Biotechnol. 14 (1996) 315.
[112] J.-H. Zhang, G. Dawes, W.P. Stemmer, Proceedings of the National Academy of Sciences 94 (1997) 4504.
[113] S.-Z. Pang, D.L. DeBoer, Y. Wan, G. Ye, J.G. Layton, M.K. Neher, C.L. Armstrong, J.E. Fry, M.A. Hinchee, M.E. Fromm, Plant Physiol. 112 (1996) 893.
[114] L. Giver, A. Gershenson, P.-O. Freskgard, F.H. Arnold, Proceedings of the National Academy of Sciences 95 (1998) 12809.
[115] K. Miyazaki, P.L. Wintrode, R.A. Grayling, D.N. Rubingh, F.H. Arnold, J. Mol. Biol. 297 (2000) 1015.
[116] P.L. Wintrode, K. Miyazaki, F.H. Arnold, J. Biol. Chem. 275 (2000) 31635.
[117] H.J. Roh, P. Kim, Y.C. Park, J.H. Choi, Biotechnol. Appl. Biochem. 31 (2000) 1.
[118] C. Vieille, G.J. Zeikus, Microbiol. Mol. Biol. Rev. 65 (2001) 1.
[119] P. Kim, Appl. Microbiol. Biotechnol. 65 (2004) 243.
[120] D.K. Oh, H.J. Oh, H.J. Kim, J. Cheon, P. Kim, J. Mol. Catal. B-Enzym. 43 (2006) 108.
[121] Y.-S. Tian, R.-H. Peng, J. Xu, W. Zhao, F. Gao, X.-Y. Fu, A.-S. Xiong, Q.-H. Yao, World J. Microb. Biot. 26 (2010) 903.
[122] H. Zhao, F.H. Arnold, Nucleic Acids Res. 25 (1997) 1307.
[123] A. Crameri, G. Dawes, E. Rodriguez Jr, S. Silver, W. Stemmer, Nat. Biotechnol. 15 (1997) 436.
[124] D.W. Leung, E. Chen, D.V. Goeddel, Technique 1 (1989) 11.
[125] K.A. Eckert, T.A. Kunkel, Nucleic Acids Res. 18 (1990) 3739.
[126] C. Neylon, Nucleic Acids Res. 32 (2004) 1448.
[127] L. Pritchard, D. Corne, D. Kell, J. Rowland, M. Winson, J. Theor. Biol. 234 (2005) 497.
[128] J. Khurana, R. Singh, J. Kaur, Mol. Biol. Rep. 38 (2011) 2919.