Biotechnology Advances 32 (2014) 308–315

Contents lists available at ScienceDirect

Biotechnology Advances

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

Research review paper

Engineering proteins for thermostability through rigidifying flexible sites

Haoran Yu, He Huang *

Department of Biochemical Engineering, School of Chemical Engineering and Technology, Key Laboratory of Systems Bioengineering, Ministry of Education, Tianjin University, Tianjin 300072, China

ARTICLE INFO

Article history:
Received 19 April 2013
Received in revised form 4 September 2013
Accepted 29 October 2013
Available online 6 November 2013

Keywords:
Protein engineering
Biocatalyst
Thermostability
Flexible sites
Rational design

ABSTRACT

Engineering proteins for thermostability is an exciting and challenging field since it is critical for broadening the industrial use of recombinant proteins. Thermostability of proteins arises from the simultaneous effect of several forces such as hydrophobic interactions, disulfide bonds, salt bridges and hydrogen bonds. All of these interactions lead to decreased flexibility of polypeptide chain. Structural studies of mesophilic and thermophilic proteins showed that the latter need more rigid structures to compensate for increased thermal fluctuations. Hence flexibility can be an indicator to pinpoint weak spots for enhancing thermostability of enzymes. A strategy has been proven effective in enhancing proteins' thermostability with two steps: predict flexible sites of proteins firstly and then rigidify these sites. We refer to this approach as rigidify flexible sites (RFS) and give an overview of such a method through summarizing the methods to predict flexibility of a protein, the methods to rigidify residues with high flexibility and successful cases regarding enhancing thermostability of proteins using RFS.

Contents

1. Introduction

Biocatalysts are increasingly used in industrial fields because of their inherent stereo-selectivity as well as their potential as a greener alternative to chemical catalysts. However, most of biocatalysts cannot tolerate high temperature since enzymes naturally evolve to work in the cellular environment. High temperature is usually required to enhance reaction rate, reactant solubility, and at the same time decrease the risk of microbial contamination in industrial field. Hence the development of enzymes with higher thermostability will increase applications of biocatalysts in industrial productions.

Various techniques have been applied to enhance thermostability of enzymes including protein engineering, posttranslational enzymatic or chemical modification, additives and immobilization. The methods of protein engineering can be grouped into directed evolution and rational design, which can be combined to semi-rational design. Directed evolution is a process using random gene mutagenesis, expression, and screening/selection to acquire stable proteins. The method has

* Corresponding author. Fax: +86 22 2740 9598.
E-mail address: huang@tju.edu.cn (H. Huang).

0734-9750/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.biotechadv.2013.10.012

been proven to be a powerful tool to improve thermostability of proteins by accumulating multiple mutations (Koksharov and Ugarova, 2011; Turner, 2009; Verma et al., 2012). Usually several rounds of evolution have to be applied and large numbers of mutants must generally be screened to obtain a desired level of change. Thus, the process of directed evolution is time and labor consuming and requires a high-throughput screening methodology that allows identification of the desired property under relevant conditions. However, not all enzyme stabilities are amenable to developing a high-throughput screening method, nor are all screening methods easy to implement at the required scale (Barrozo et al., 2012). In rational design biochemical data, protein structures and molecular modeling data are evaluated to propose mutations that would be introduced by site-specific mutagenesis (Anbar et al., 2012; Imani et al., 2010; Watanabe et al., 2006). Rational design methods are universal, faster and have the potential to be developed into algorithms that can quantitatively predict the stabilities of the designed sequences (Schweiker and Makhatadze, 2009). Nevertheless, despite the availability of a fast-growing number of protein structures and sophisticated computational algorithms, pure rational design is still limited by some factors such as incomplete understanding of structural features and their contributions to function and limited understanding of protein dynamics (Ruscio et al., 2009; Steiner and Schwab, 2012a). Semi-rational design uses information derived from structural data to identify amino acids in interesting regions and then mutated them randomly or by site-saturation mutagenesis one by one or in combination (Chica et al., 2005). This approach combines the advantages of rational and random protein design creating smaller smarter libraries and makes the directed evolution faster and more efficient (Blum et al., 2012; Reetz et al., 2010a).

Recently, a protein engineering approach, rigidify flexible sites has been demonstrated to be pronouncedly effective in increasing thermostability of proteins. System structural analysis of individual proteins from mesophilic and thermophilic organisms with known three-dimensional structure has shown that the latter are characterized by higher degrees of rigidity caused by increasing strength of interaction between contact residues, salt-bridge strength, better packing of hydrophobic interactions and better developed hydrogen bond network (Mamonova et al., 2013; Reetz et al., 2006). Similarly, comparison analysis of structure properties between psychrophilic and mesophilic proteins revealed that psychrophilic enzymes are also more flexible (Paredes et al., 2011). Therefore, it is possible to enhance the thermostability of enzymes through increasing the rigidity at appropriate sites. Highly flexible residues have a low number of contacts with other amino acids and produce a local perturbation inside the complex network of non-covalent connections. They can trigger protein unfolding due to their large thermal fluctuations. Introducing mutations to rigidify these residues would be an effective way to improve stability. Hence, flexibility can be applied as an indicator to find out weak spots. Once the weak spots were identified, further thermostability of protein could be achieved by optimizing weak spot regions through methods such as saturation mutagenesis, site-directed mutagenesis of these regions or neighboring residues.

The aim of this review is to provide an introduction to the protein engineering approach to improve thermostability of proteins: rigidify flexible sites. We name it as RFS. Table 1 summarizes recent successful cases achieved by RFS. Some of them are discussed in detail while others are provided in the table to give a general summary of the work and to direct interested readers toward the references. Since we mainly focus on engineering proteins for thermostability other studies using RFS to improve other properties of enzymes such as stability toward hostile organic solvents are not mentioned here (Reetz et al., 2010b). The method, RFS, contains two steps: predict flexible sites and rigidify these sites. There are many approaches available for predicting residues with high flexibility and rigidifying flexible sites and these methods will also be discussed in detail below. To our knowledge we are the first to attempt to give an overview of such a method.

2. Computational approaches to predict flexible sites

Protein flexibility can be monitored by experimental approaches or computer-aided methods. Some of experimental approaches have been reviewed by Kamerzell and Middaugh (2008) that include nuclear magnetic resonance spectroscopy, hydrogen isotope exchange, high resolution ultrasonic spectroscopy pressure perturbation calorimetry (PPC) and so forth. Here three commonly used methods in RFS are discussed in detail and other computer-aided methods will be simply mentioned.

2.1. B-FITTER

B-factor (or B value) is commonly used to represent flexibility. B-factor indicates atomic displacement parameters obtained from X-ray data that reflect searing of atomic electron densities with respect to their equilibrium position due to positional disorder or thermal motion (Parthasarathy and Murthy, 2000). The residues with higher B factor are more flexible and more thermo-unstable. The average B-factors for a residue can be calculated by averaging the B-values of each atom in the amino acid with a program, B-FITTER (Reetz and Carballeira, 2007). Access to B-FITTER is shown in Table 2. Using this approach to predict flexible residues requires availability of X-ray structure of wild type enzymes. B-values vary greatly from protein to protein due to differences in crystal quality and structural refinement. In order to acquire best results, a few criteria have been mentioned to choose the PDB files including a high-resolution structure, a single-crystal structure, a structure without any external molecule bound and a reduced number of non-resolved atoms (Reetz and Carballeira, 2007). However, several methods have been developed to compute B-factor in the cases that crystallographic data is not available (Schlessinger, 2006; Yuan et al., 2005). Therefore, it is possible to apply RFS to stabilize enzymes when their X-ray structures are not available. B-FITTER has been proven effective in predicting flexible sites and successfully used to increase thermostability of enzymes. Nevertheless, some limitations should be realized when applying this method to predict flexible regions. Protein flexibility in solution could differ qualitatively from that in a crystal. Fluctuations tend to be larger in solution than the flexibility estimated by B-factors in a crystal especially in the flexible loop regions. It is reasonable to use B-factors to infer atomic fluctuations for backbone atoms, whereas it will underestimate the degree of motion for flexible side chains (Eastman et al., 1999). One possible reason for this difference may be the assumption made in calculating B-factors that the fluctuations are harmonic. This assumption has been shown to result in an underestimation of fluctuation magnitudes (Garcia et al., 1997). Another possible contribution is the presence of crystallization agents, ligands or large-size ions in the experiment. These materials could restrict protein motion due to steric hindrance or salt bridges (Meinhold and Smith, 2005). In addition, X-ray diffraction pattern arises from the molecule in solid state, and crystal-packing interactions can introduce artifacts. If the crystal-packing effects are considered, the B-factor is in better agreement with flexibility in solution (Song and Jernigan, 2007). Furthermore, crystallographic data is usually recorded at temperature around 110 K and the molecules are frozen in a conformation that does not necessarily represent the physiological state (Teilum et al., 2009).

2.2. Molecular dynamics simulations

MD (molecular dynamics) simulations are another widely used method to predict flexibility of proteins. Different from B-FITTER, MD simulations focus on the flexible motions of proteins during a period and provide information on the motional properties of atoms in a protein structure. By performing simulations at different temperature or by running nonequilibrium simulations, motional properties of the structure during thermal unfolding can be analyzed and unfolding regions or weak spots can be predicted (Radestock and Gohlke, 2011).