Biochimica et Biophysica Acta 1854 (2015) 10–19

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Clostridium thermocellum thermostable lichenase with circular permutations and modifications in the N-terminal region retains its activity and thermostability

A.A. Tyurin [a,b], N.S. Sadovskaya [a], Kh.R. Nikiforova [a], O.N. Mustafaev [a], R.A. Komakhin [c], V.S. Fadeev [a], I.V. Goldenkova-Pavlova [a,b,*]

[a] Institute of Plant Physiology, Russian Academy of Sciences, ul. Botanicheskaya 35, Moscow 127276, Russia
[b] Department of Genetics and Biotechnology, Russian State Agrarian University—Moscow Timiryazev Agricultural Academy, ul. Timiryazevskaya 49, Moscow 127550, Russia
[c] All-Russia Research Institute of Agricultural Biotechnology, Russian Academy of Agricultural Sciences, ul. Timiryazevskaya 42, Moscow 127550, Russia

ARTICLE INFO

Article history:
Received 26 June 2014
Received in revised form 25 September 2014
Accepted 15 October 2014
Available online 23 October 2014

Keywords:
Lichenase
Clostridium thermocellum
Circular permutation
Carrier protein
Thermostability
Activity

ABSTRACT

The *Clostridium thermocellum* lichenase (endo-β-1,3;1,4-glucan-D-glycosyl hydrolase) displays a high thermostability and specific activity and has a compact protein molecule, which makes it attractive, in particular, for protein engineering. We have utilized *in silico* analysis to construct circularly permuted (CP) variants and estimated the retained activity and thermostability. New open termini in the region of residues 53 or 99 in two lichenase CP variants (CN-53 and CN-99) had no effect on their activity and thermal tolerance versus another variant CP variant, CN-140 (cut in the region of residue 140), which displayed a dramatic decrease in the activity and thermostability. Construction and further activity and thermostability testing of the modified lichenase variants (M variants) and CP variants with peptides integrated via insertion fusion have demonstrated that the N-terminal regions in the lichenase catalytic domain (53 and 99 amino acid residues) that permit circular permutations with retention of activity and thermostability of the enzyme as well as the region between the C and N termini of the native lichenase in thermostable and active lichenase variants (CN-53 and CN-99) may be used for integrating small peptides without the loss of activity and thermostability. These findings not only suggest that CP predictions can be used in search for internal integration sites within protein molecule, but also form the background for further enzymatic engineering of the *C. thermocellum* thermostable lichenase aiming to create new fusion proteins.

1. Introduction

Protein engineering is targeted to the creation of the molecules with necessary specified properties. Thermostable proteins are of special interest for protein engineering. The efforts here are directed to both the search for new thermostable proteins and modification of the already known ones. There are various motivations in the engineering of thermostable proteins, including study of the functional role of protein structural elements, determination of the folding, decrease in the proteolytic susceptibility, improvement of catalytic activity and/or thermostability, and search for internal sites allowing for integration of other proteins [1–5].

The following approaches are typically used in protein engineering: deletions (excision of domains and/or individual amino acids); random and directed mutagenesis; grouping of two or more enzyme catalytic modules; and circular permutation (CP) method [4–9]. Note that the CP method is a powerful molecular tool for altering protein sequences and, as a consequence, their structure [9,10]. In particular, "cutting" a protein polypeptide chain within a functionally significant element may change the conformation of this protein, thereby leading to loss of its function; however, this "cutting" beyond the functional element or on its boundary frequently causes only insignificant changes in the protein function [11,12]. In other words, the properties of a "cut" protein will vary depending on the location of this breakpoint. Based on this strategy, protein molecules are "cut" to construct and characterize circularly permutated (CP) proteins, detecting the protein functional elements important for their properties via a comparative analysis [9].

An end-to-end fusion is frequently used when engineering bifunctional proteins to connect two enzymes [8,13–16] using different approaches, for example, by optimization of peptide linkers [14] or through all-atom molecular dynamics simulations [16]. However, the chimeric enzymes constructed with the help of this method are frequently more susceptible to proteolytic degradation and structural instability [17]. In such cases, the strategy for constructing chimeric enzymes using the "insertion" fusion when one gene is inserted within the other gene may be more beneficial [12,18,19]. The potential regions

* Corresponding author at: Institute of Plant Physiology, Russian Academy of Sciences, ul. Botanicheskaya 35, Moscow,127276, Russia. Tel.: +7 499 231 8315.
E-mail address: irengold58@gmail.com (I.V. Goldenkova-Pavlova).

http://dx.doi.org/10.1016/j.bbapap.2014.10.012
1570-9639/

permitting integration are of a paramount importance for successful construction of such fusion proteins. We have assumed that the regions permitting circular permutations in a protein molecule, that is, the regions where new N and C termini of the protein are formed without any significant loss in the major protein properties, may be potentially used also as the regions for insertion fusions with protein sequences.

Glycosyl hydrolases are used as models when studying the folding of protein molecules and the mechanisms rendering them thermostable [20,21]. In addition, a practical application has been found for glycosyl hydrolases [13,22], in particular, for constructing bifunctional proteins [14,19].

In this work, the *Clostridium thermocellum* lichenase (endo-β-1,3;1,4-glucan-D-glycosyl hydrolase, EC 3.2.1.73, P29716) was used as a model. Similar to the majority of glycosyl hydrolases, lichenase is a single subunit protein comprising (i) a signaling peptide (positions 1–27); (ii) a catalytic module, homologous to the full-sized *Bacillus* lichenase (positions 28–251); (iii) a Thr–Pro box (positions 252–261); and (iv) a cellulosome-binding (dockerin) domain (positions 262–334) [22]. It has been earlier demonstrated that the lichenase deletion variant carrying the catalytic domain and part of Thr–Pro box as well as the lichenase catalytic domain alone retains a high specific activity and thermostability [23]. A compact structure of the lichenase molecule, its high thermostability, and specific activity are attractive, in particular, for constructing hybrid proteins. However, it is necessary to clarify which particular regions in the lichenase catalytic domain permit internal integration without a dramatic loss in its activity and thermostability. Note that the precise information about the role of structural elements (modules) constituting the lichenase catalytic domain in manifestation of its major enzymatic properties is currently rather limited [7,24].

In this work, we have utilized *in silico* analysis of the lichenase sequences, in particular, by predicting potential circular permutation sites, to construct three CP lichenase variants and experimentally determine their properties (activity and thermostability). New open ends in the N-terminal region (amino acid residues 53 or 99), unlike the open ends in the C-terminal region (residue 140), had no effect on the activity or thermal tolerance of the lichenase CP variants. We have constructed and tested activity and thermostability of the modified lichenase variants (M variants) with peptide insertion fusions and found out that the CP regions are also suitable for integrating peptides. It has been clearly demonstrated that the N-terminal regions of lichenase catalytic domain that permit CP with retention of the enzyme activity and thermostability also permit internal integration of small peptides, including multiple integration. Moreover, it has been shown that the lichenase CP variants that retained their activity and thermostability also permit insertion fusions with small peptides as internal modules between the C and N termini of the native lichenase without any dramatic decrease in their thermostability and activity. As a result of this study, potential sites that can be used to integrate other proteins as internal modules have been detected in the lichenase and its CP variants, namely, three variants of the catalytic domains and two variants of the CP variants of this enzyme carrying the integration sites and retaining high thermostability and activity are proposed, which have a potential for further designing of bifunctional proteins.

2. Materials and methods

2.1. In silico analysis

Homologs of thermostable lichenase were identified using the *Blast* software [25]. BLAST was performed to search the potential templates available in the PDB database [26]. The homology model of lichenase was built with *Phyre* [27], allowing for 3D structure prediction for the proteins with a known PDB structure used as a template pattern. Of all the proteins that *Phyre* suggested as a pattern, we selected the 3I4I protein as displaying the maximal homology to LicBM3 and obtained the 3D structure for lichenase. The program *Muscle* [28] was used to construct pairwise alignments of amino acid sequences. The circularly permutated lichenase homologs were searched for using *iSARST* web server [29] and the *Blast* software in PDB [26]. Potential CP sites in proteins were predicted using the *CPred* web server [30].

2.2. Recombinant DNA techniques

Standard molecular cloning procedures were used in the work as well as Evrogen (Russia) primers and Promega (United States), Fermentas (Lithuania), QIAexpress (United States), and Novagene (United States) enzymes and reagent kits. Table S1 lists the primers used in cloning procedures. The sequence of *lic*BM3 deletion mutant was determined by PCR with the plasmid pQE-*lic*BM2-KM2-Mys25 [31] as a template and the primers L-Forw and L-Rev. The *Bam*HI–*Pst*I PCR fragment was cloned into the plasmid pQE30 (Qiagen, United States) hydrolyzed with the *Bam*HI and *Pst*I restriction endonucleases to get the plasmid pQE-*lic*BM3.

2.2.1. Constructing the hybrid genes encoding circularly permutated (CP) lichenase variants

Initially, *lic*BM3 gene fragments, designated C53, C99, and C140, were produced by PCR using pQE-*lic*BM3 as a template and the primers C53-For/C-Rev, C99-For/C-Rev, and C140-For/C-Rev, respectively, as well as *lic*BM3 gene fragments, designated as N53, N99, and N140, using the primers N-For/N53-Rev, N-For/N99-Rev, and N1-For/N140-Rev, respectively. Synthesized PCR fragments C53, C99, and C140 were hydrolyzed with *Bam*HI restriction endonuclease and PCR fragments N53, N99, and N140, with *Pst*I. Then the fragment pairs *Bam*HI-C53/*Pst*I-N53, *Bam*HI-C99/*Pst*I-N99, and *Bam*HI-C140/*Pst*I-N140 were ligated to clone each pair into the pQE30 hydrolyzed with *Bam*HI and *Pst*I. The resulting expression vectors were designated pQE-CN-53, pQE-CN-99, and pQE-CN-140.

2.2.2. Constructing the hybrid genes encoding the lichenase catalytic module with integrated small peptides

The plasmid pQE-NC-L-53 was constructed in several stages. The *lic*BM3 gene fragment designed N-53-1 was obtained by PCR with pQE-*lic*BM3 as a template and F1/R1-53 primer pair. At the next stage, the PCR fragment designated N-53 was produced using N-53-1 PCR fragment as a template and F1/R2-53 primer pair and the pQE-*lic*BM3 gene fragment designated C-53 using pQE-*lic*BM3 as a template and F-53/R1 primer pair. N-53 and C-53 PCR fragments were hydrolyzed with *Apa*I restriction endonuclease and ligated. The ligation product was further used as a template for PCR with F1 and R1 primers. The PCR fragment designated NC-L-53 was hydrolyzed with *Sph*I and *Pst*I and cloned into the plasmid pQE30 hydrolyzed with *Sph*I and *Pst*I.

The plasmid pQE-NC-L-53-99 was also constructed in several stages. The *lic*BM3 gene fragment designated N-99-1 was produced by PCR using pQE-NC-L-53 as a template and F1/R1-99 primer pair. Then, the PCR fragment designated N-99 was obtained with N-99-1 PCR fragment as a template and F1/R2-99 primer pair and the *lic*BM3 gene fragment designated C-99 was obtained using pQE-*lic*BM3 as a template and F-99/R1 primer pair. N-99 and C-99 PCR fragments were hydrolyzed with *Sma*I and ligated. The ligation product was used as a template for PCR with F1 and R1 primers. The resulting PCR fragment, designated NC-L-53-99, was hydrolyzed with *Sph*I and *Pst*I and cloned into the pQE30 hydrolyzed with the same restriction endonucleases.

The plasmid pQE-NC-L-99 was constructed by cloning *Apa*I–*Pst*I fragment of the plasmid pQE-NC-L-53 into the pQE-NC-L-53-99 hydrolyzed with *Apa*I and *Pst*I restriction endonucleases.

The plasmid pQE-NC-L-140 was constructed in several stages. The *lic*BM3 fragment designated N-140-1 was produced by PCR using pQE-*lic*BM3 as a template and F1/R1-140 primer pair. Then the PCR fragment designated N-140 was synthesized with N-140-1 PCR fragment as a template and F1/R2-140 primer pair and the *lic*BM3 gene