BBA - Proteins and Proteomics 1865 (2017) 1105–1113

Contents lists available at ScienceDirect

BBA - Proteins and Proteomics

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

A glimpse into the specialization history of the lipases/acyltransferases family of CpLIP2

Anne-Hélène Jan[a], Eric Dubreucq[a], Jullien Drone[a,b], Maeva Subileau[a,*]

[a] Montpellier Supagro, UMR 1208 IATE, - 2 place Viala, 34060 Montpellier cedex 2, France
[b] Ecole Nationale Supérieure de Chimie de Montpellier, 8 rue de l'Ecole Normale, 34296 Montpellier, France

ARTICLE INFO

Keywords:
Ancestral protein resurrection
Lipase
Acyltransferase
Structure/function

ABSTRACT

Lipases/acyltransferases homologous to CpLIP2 from *Candida parapsilosis* belong to the α/β hydrolase superfamily as lipase A from *Moesziomyces antarcticus* (*Candida antarctica*), and constitute a consistent phylogenetic subgroup with at least 56% identity. Lipases/acyltransferases share the phenotypic characteristic of a high acyltransfer activity even in aqueous media with very high water thermodynamic activity. Previous mutagenesis and evolution strategies have given insights into the role of key residues and protein subdomains in the reaction and substrate specificities of these enzymes. However, multiple mutations are often deleterious for the activity and the identification of all the residues that historically led to the function is complicated. A new complementary approach to elucidate structural determinant was conducted in this study, based on the resurrection of ancestral proteins to understand how the evolution led to the present properties of the biocatalysts. By doing so, the comparison with the extant proteins can lead to the identification of key residues involved in the enzymes' specialization. Using Ancestral Sequence Reconstruction, we have generated a putative ancestral lipases/acyltransferases, PaleoLAc. This enzyme shares a high level of identity with CpLIP2 but has a different catalytic behavior. PaleoLAc allowed the identification of putative key residues involved in acyltransfer ability and supports the hypothesis that this exceptional property within the lipases/acyltransferases family is linked to a cluster of residues in the vicinity of the active site. As a representative of the ancestral origin of the diversity of the catalytic behaviors observed in modern lipases/acyltransferases, PaleoLAc constitutes a powerful tool for further engineering toward targeted specialization.

1. Introduction

Within the abH38 α/β hydrolase superfamily of the Lipase Engineering Database (LED) [1], represented by lipase A from *Moesziomyces antarcticus* (*Candida antarctica*) (CAL-A), a peculiar family that constitutes a consistent phylogenetic subgroup (> 56% of identity in their primary sequence) was evidenced: the lipases/acyltransferases family [2,3]. They are predicted to share with CAL-A its structural peculiarities [2,4]: globular class Y α/β serine hydrolase of about 450 amino-acids exhibiting, in addition to the main core of the protein, two distinct sub-domains called the "cap" and the "flap" [2,5–7]. In comparison to other α/β hydrolases, CAL-A and its homologs have very divergent sequences, particularly in the cap region which is exceptionally long (about 100 residues) [2,7]. The cap has been shown to play an important role for the acyl moiety specificity [8] and may comprise a hydrophobic tunnel-like binding site comparable to that of the *Candida rugosa* lipases (CRL) [2]. The structure and role of the C-terminal flap is still unclear but that of CAL-A was proposed to play a "lid-like" role in the interfacial activation of the enzyme [6], and was also compared to the lid in CRL [2].

The lipases/acyltransferases share the unique ability to catalyze preferentially alcoholysis over hydrolysis when a suitable nucleophile is present, even in media with high water content and high thermodynamic activity of water a_W > 0.95 [9]. Indeed like some penicillin acylases or proteases [10–14], under "kinetically controlled reaction", these lipases/acyltransferases favor the acyltransfer reaction and allow achieving transient product yield superior to that of the thermodynamic equilibrium [15]. In comparison, with lipases such as the lipase B from *M. antarcticus* (*C. antarctica*) CAL-B, the maximum yield is reached under "thermodynamically controlled reaction" [15] and high product concentration can only be obtained in quasi-anhydrous conditions, that often require a tight regulation of a_W to prevent competitive hydrolysis while maintaining an optimal enzyme hydration [16–18]. Over more than 20 years, several members of the lipases/acyltransferases family have been identified and characterized in *Candida parapsilosis* (CpLIP2), *Candida albicans* (CaLIP4, CalLAc5 and CalLAc8), *Candida tropicalis*

* Corresponding author.
E-mail addresses: maeva.subileau@supagro.fr, subileau@supagro.inra.fr (M. Subileau).

http://dx.doi.org/10.1016/j.bbapap.2017.06.004
Received 4 January 2017; Received in revised form 6 June 2017; Accepted 8 June 2017
Available online 13 June 2017
1570-9639/

(CtroL4) and *Candida dubliniensis* (CduLAc) [3,9,15,19–24]. Recently, based on the studies of Kasche [12], Subileau et al. proposed a reaction scheme and simple methodology to assess the acyltransferase potential of lipases [15] ranging from enzymes exhibiting low (hydrolases) to medium and high acyltransfer ability (acyltransferases). According to this classification, maximum product concentration superior to that of the thermodynamic equilibrium can only be obtained with the so-called "acyltransferases". The characterized enzymes of the lipases/acyltransferases phylogenetic group exhibit medium to high acyltransferase ability and display various substrate specificities (medium or long chain selectivity for the acyl-donor substrate and preference for unbranched primary alcohols for the acyl-acceptor alcohol). Among the characterized enzymes of the family, CpLIP2, CtroL4, CaLIP4 and CalLAc8 are cold active enzymes, meaning that they exhibit high catalytic activity at near-zero to room temperatures [7,23,25].

To further understand the structure function relationships involved in the acyltransfer ability and substrate specificities of these unique biocatalysts, the first approach was to use site directed mutagenesis. Structure comparison between CAL-A and homology 3D models of CpLIP2 and of CduLAc enlightened differences in size and hydrophobicity of residues in the vicinity of the active site [4,24]. Müller et al. [26] and Jan et al. [24] have turned CAL-A into a lipase/acyltransferase with mutations increasing the hydrophobicity in this region. Another approach was to generate chimeric enzymes (rational gene shuffling) with exchanges of subdomains to assess the role of the cap and the C-term flap in the enzyme reactivity. By doing so, Jan et al. [7] showed the implication of the cap in the substrate specificity and suggested the preponderant role of the main core of the protein in the transesterification ability.

Despite these discoveries, it is difficult to precisely identify all the residues involved in acyltransfer or substrate specificity [27,28]. As a matter of fact, in the lineage of the lipases/acyltransferase, the sequential evolution that led to their peculiar catalytic behavior is based on multiple mutations of interacting residues inside the gene sequence. As many neutral and epistatic mutations may have accumulated during the evolution of the proteins under comparison, the exchange of putative key-residues by site-directed mutagenesis can lead to non-functional proteins [29,30]. The identification of residues that, in an evolutionary perspective, have led to the new function is often impossible by only comparing the extant sequences [29]. Complementary to the results obtained with the rational design strategies based on sequences and structure comparisons of lipases/acyltransferases, another approach is to design a putative ancestral protein of the family [31,32]. The ancestral sequence reconstruction (ASR) is the calculation of putative ancient protein sequences on the basis of the extant ones. Based on primary sequences, corresponding alignment and phylogenetic tree of the proteins of interest, this technique generates a statistical model for ancestral primary sequences at each node of the phylogenetic tree. Each amino acid of this new sequence is evaluated in terms of likelihood. ASR was used in numerous evolutionary studies, such as the evolution of ancestral archosaur visual pigment rhodopsin [33] or the evolution of GFP-like proteins [34–36]. ASR was also applied in protein engineering, including the design of more thermostable proteins [37]. Interestingly, because the most probable residue is selected at each site, many ASR projects presented higher thermostability than the extant proteins [38]. Researches on the structural determinants involved in the proteins functions, as the distinct ligand specificities of the clades of estrogen receptors and non-aromatized steroid receptors [39] or the product specificity of terpene synthases [40] were also carried on.

Thus, we used this different approach on the lipases/acyltransferases family, hoping to elucidate some of the key residues responsible for their unique catalytic behavior. We first designed one putative ancestral protein of the family using FastML program [41]. Then, after its production, we characterized its acyltransfer ability and its substrate specificity. We enlightened these results with a comparative study of the structure and sequence of the other characterized members of the family (CpLIP2, CtroL4, CduLAc, CalLAc8, CalLAc5). The lipase CAL-A was used as an external reference in this study. It also appeared particularly interesting because CAL-A catalytic behavior is significantly different from the lipases/acyltransferases: its acyltransfer ability is low, its substrate specificity is centered on C14, and it exhibits exceptional thermo-activity [3,7,25].

2. Experimental

2.1. Phylogenetic analysis and ancestral resurrection

The sequences of lipases/acyltransferases were collected from the national Center of Biotechonlogy Information (NCBI) [42] using protein alignement BLAST [43,44] versus the non-redundant protein sequence database. Sequence visualization and multisequence alignments were generated using Seaview [45] and ClustalX [43,46]. The alignment was subsequently submitted to PhyML 3.0 software [47] with the LG substitution matrix with 100 iterations to generate a phylogenetic tree. The prediction of the putative ancestral proteins at each node was performed using FastML [41]. We have chosen the putative ancestor at the root of the family and called it PaleoLAc.

2.2. Construction and analysis of the 3D model

The 3D model of CpLIP2 was previously described by Subileau et al. [4]. The 3D model of PaleoLAc was designed, based on the crystallographic structure of CAL-A (PDB ID: 2VEO, 3GUU) [5], using Modeller 9.14 [48–50] via UCSF Chimera 1.10.2 [51]. The cavities inside the proteins were determined using the online program CASTp [52].

2.3. Plasmids, strains and reagents

Competent cells of *Escherichia coli* XL1-Blue MRF' (Stratagene la Jolla, CA, USA) were used for DNA propagation. *Komagataella pastoris* (*Pichia pastoris*) X-33, pPICZαB vector and zeocin were from Invitrogen (Life Technologies SAS, Saint Aubin, France). SacI restriction enzyme and RNaseA were purchased from Roche Diagnostic AG (Rotkreuz, Switzerland). All lipase substrates and reagents were purchased from Sigma-Aldrich (Lyon, France) and were of analytical grade.

2.4. Lipase expression and production

Protein sequence of the lipases acyltransferases CduLAc from *C. dubliniensis* [24], CalLAc5, CalLAc8 and CaLIP4 from *C. albicans* [23], and CtroL4 from *C. tropicalis* [3]and of the lipase CAL-A from *M. antarctica* were obtained from NCBI databases under accession numbers XP_002421466.1, XP_717001.1,XP_711685, AF191317, XP 002548755 and 2VEO_A respectively. The sequence of CpLIP2 from *C. parapsilosis*, accessible under number CAC86400.1, was from Neugnot et al. [22], with the signal peptide suppressed. The protein sequence of PaleoLAc was converted in nucleotid one using optimized codons for *K. pastoris*. DNA was synthesized and subcloned by Life Technologies (Regensburg, Germany) in pPICZαB in fusion with the signal peptide of the alpha-mating factor of *Saccharomyces cerevisiae* provided in the plasmid, as described by Neang et al. [3]. Clones for the expression of PaleoLAc were produced as described by Neang et al. [3].

2.5. Heterologous production in bioreactor

Recombinant lipases were obtained from culture supernatants of transformed *K. pastoris* as described by Brunel et al. [53] and modified by Neang et al. [3]. The enzyme extract preparation was executed as described by Jan et al. [24].