FEBS Open Bio 4 (2014) 121–127

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

# Ancestral mutations as a tool for solubilizing proteins: The case of a hydrophobic phosphate-binding protein☆

Daniel Gonzalez[a], Julien Hiblot[a], Nune Darbinian[b], Jernelle C. Miller[b,c], Guillaume Gotthard[a], Shohreh Amini[b,c], Eric Chabriere[a,*], Mikael Elias[d,*]

[a] URMITE UMR CNRS-IRD 6236, IFR48, Faculté de Médecine et de Pharmacie, Université de la Méditerranée, Marseille, France
[b] Department of Neuroscience, Temple University School of Medicine, Philadelphia, PA 19140, USA
[c] Department of Biology, College of Science and Technology, Temple University, Philadelphia, PA 19122, USA
[d] Weizmann Institute of Science, Biological Chemistry, Rehovot, Israel

ARTICLE INFO

*Article history:*
Received 26 November 2013
Received in revised form 21 December 2013
Accepted 23 December 2013

*Keywords:*
Ancestral librairies
Protein engineering
Protein solubilization
Hydrophobic proteins
Phosphate-binding proteins
DING proteins

ABSTRACT

**Stable and soluble proteins are ideal candidates for functional and structural studies. Unfortunately, some proteins or enzymes can be difficult to isolate, being sometimes poorly expressed in heterologous systems, insoluble and/or unstable. Numerous methods have been developed to address these issues, from the screening of various expression systems to the modification of the target protein itself. Here we use a hydrophobic, aggregation-prone, phosphate-binding protein (HPBP) as a case study. We describe a simple and fast method that selectively uses ancestral mutations to generate a soluble, stable and functional variant of the target protein, here named sHPBP. This variant is highly expressed in *Escherichia coli*, is easily purified and its structure was solved at much higher resolution than its wild-type progenitor (1.3 versus 1.9 Å, respectively).**



## Introduction

Stable, soluble and functional proteins comprise ideal models for functional and structural studies. However, when overexpressed in heterologous systems such as in *Escherichia coli*, natural proteins from various sources can sometimes be insoluble, unstable or poorly expressed [1]. These difficulties considerably hamper studies of certain proteins, and have yielded a considerable bias in protein functional and structural analysis toward soluble and expressible proteins [2]. Several strategies have been developed to skirt these limitations. Classical methods involving expression in heterologous systems usually screen the host nature, the culture conditions, and media composition [1,3–5]. Codon optimization, protein fusion or the co-expression with chaperones [6] may also represent useful strategies to successfully express proteins [1,5,6]. Nevertheless, these trials may remain inefficient in some cases; particularly for numerous mammalian proteins [1]. Thus, methodologies tuning the protein target itself emerged, with the aim of producing soluble and expressible models for further studies [7,8]. In that respect, site-directed mutagenesis could be used to substitute surface residues and therefore generate more soluble proteins. However, this technique is limited by our ability to precisely identify problematic surface residues [9]. Conversely, directed evolution allows to extensively mutate the target protein and to select for more soluble variants [10]. The explored sequence space is however huge, and the method thus requires a high throughput screening method [11]. An alternative method, called DNA shuffling, uses several genes sharing high sequence identity with the target protein and shuffles them all. The screening of the resulting gene library for solubility may yield soluble and expressible variants, which can subsequently be subjected to functional and structural studies [7].

In regard to protein stabilization or solubilization, phylogenetic-based protein engineering may represent a powerful method. Indeed, consensus libraries, which are composed of mutations that bring the sequence of the target gene closer to the family consensus sequence, can efficiently stabilize proteins [12]. Ancestral mutations also have proven ability to yield soluble and stabilized protein variants [13,14]. Moreover, the use of ancestral mutations libraries can yield interesting protein variants with altered enzymatic activity and/or stability [11,15–18]. Ancestral mutations may therefore be turned into an easy and fast method to solubilize/stabilize contemporary proteins.

In this study, we focus on ancestral mutations, and their ability to solubilize a protein target. We therefore used the human phosphate binding protein (HPBP) as a case study. HPBP belongs to DING proteins



* Corresponding authors. Tel.: +33 4 91 32 43 75; fax: +33 4 91 38 77 72. Tel.: +972 0 8 934 6100; fax: +972 8 934 4118 .
*E-mail addresses:* eric.chabriere@univ-amu.fr (E. Chabriere) mikael.elias@gmx.fr (M. Elias).

2211-5463/$36.00 
http://dx.doi.org/10.1016/j.fob.2013.12.006

family, a clade of the phosphate binding protein (PBP) superfamily [19]. HPBP is a hydrophobic, possibly an apolipoprotein, crystallized from supposedly pure human paraoxonase (PON1) preparations [20–22]. HPBP possesses a venus-flytrap topology identical to the high affinity phosphate-binding protein (PBP or PstS) carriers of the ABC transporter systems, and a similar phosphate-binding ability [22]. Interestingly, and as for other related DING proteins [23–25], HPBP has been shown to inhibit HIV-1 replication by targeting the transcriptional step [26].

Nevertheless, functional and structural studies on HPBP are considerably hampered by its high hydrophobicity [22], its propensity to aggregate, and the failure to express it heterologously in soluble form (Chabriere, unpublished results). The existing purification procedure, starting from human plasma samples, is complex and laborious, and yields to little amounts of pure HPBP [20]. We hereby describe a simple and fast method to generate an ancestral-mutations based, fully functional, soluble variant of HPBP.

## Materials and methods

### *Phylogenetic analysis and ancestral resurrection*

The sequences of phosphate-binding proteins (including DING and PstS proteins) were collected from the National Center of Biotechnology Information (NCBI) using protein alignment BLAST (blastp) [27,28] with default parameters *versus* the non-redundant protein sequence database (nr). Only complete protein sequences were selected, and redundancy was subsequently removed (maximum 95% of sequence identity) with Cd-hit [29]. The sequence alignment was performed with clustalW 2.0 software [30] and manually improved (Fig. 1-1). The substitution matrix corresponding to the sequence alignment was determined using the Prottest software [31]. The alignment was subsequently submitted to PhyML software [32] with the JTT substitution matrix with 100 iterations. The prediction of the putative ancestral sequences at each nodes was performed using FastML [33]. We have chosen the putative ancestor of HPBP and a related, bacterial, soluble homologue, PfluDING from *Pseudomonas fluorescens* (Node 10) (Fig. 1-2) [34,35]. The ancestral mutation library contains the substitutions of the putative ancestral sequence, as compared to HPBP sequence. By applying four simple filters:

(1) include substitutions of surface apolar residue into polar residue,
(2) include substitution from Gly to X, with the exception of Gly residue involved in the start/end of secondary structure,
(3) include core mutations,
(4) include mutations that change a surface hydrophobic residue into a less hydrophobic one, we have reduced the numbers of substitutions from 93 to 22 (Fig. 1-3).

### *Gene synthesis and cloning of HPBP and sHPBP*

The genes encoding for HPBP and sHPBP were optimized for *E. coli* expression and synthesized by service providers (Genecust, Luxembourg, and GeneArt, Life Technologies, France, respectively) (Fig. 1-4). The genes were subsequently subcloned into pET22b ( + ) (Novagen) using *Nco*I and *Xho*I as cloning sites.

### *Production and purification of HPBP and sHPBP*

Productions of HPBP and sHPBP were performed using *E. coli* BL21(DE3)-pGro7/GroEL cells (TaKaRa) in 6 l of ZYP medium [1] (100 μg $ml^{-1}$ ampicillin, 34 μg $ml^{-1}$ chloramphenicol). The cultures were grown at 37 °C to reach $OD_{600nm}$ = 0.6 and then induced by starting the consumption of lactose in ZYP medium coupled to temperature transition to 17 °C during 16 h. Cells were harvested by centrifugation (4500*g*, 4 °C, 15 min) and pellets were suspended in 400 ml of lysis buffer (20 mM TRIS, pH 8, 100 mM NaCl, Lysozyme 0.25 mg $ml^{-1}$, DNAse I 10 μg $ml^{-1}$, PMSF 0.1 mM, $MgSO_4$ 20 mM and 8 tablets of anti-protease EDTA-free (Roche)) and stored at −80 °C for 2 h. Cells were then thawed at 37 °C for 15 min and disrupted by 3 steps of 30 s of sonication (QSonica sonicator; amplitude 40). Debris was removed by centrifugation (12,500*g*, 4°C, 30 min). Supernatant was loaded on a Nickel affinity column (HisTrap 5 ml, FFCrude from GE Healthcare) at a flow rate of 5 ml $min^{-1}$. Proteins gripped to the column were eluted by imidazol, using an elution buffer (20 mM TRIS, pH 8, 100 mM NaCl and 250 mM imidazole). Then, a size exclusion chromatography step (Superdex 75 16/60, GE Healthcare) was performed using buffer 20 mM TRIS, pH 8 and 100 mM NaCl. Protein production and purity were checked by 15% SDS–PAGE analysis (Fig. 1-5) and mass spectrometry analysis (Plateforme Timone, Marseille, France).

### *Crystallization of sHPBP*

sHPBP was concentrated to 4 mg $ml^{-1}$ using a centrifugation device (Vivaspin 500, MWCO 3 kDa, Sartorius stedim, Germany). Crystallization trials were performed at 298 K using the same condition as the homologue PfluDING (*i.e.* $LiSO_4$ 1 M, 20–30% PEG 8000 and Sodium Acetate 2 M at pH 4.5–5.5, [34,36,37]). Since only thin crystal plates were obtained, commercial screens conditions were tested (*i.e.* Stura and MDL, Molecular Dimension, England) using a Mosquito instrument (TTP Labtech, England) with the sitting-drop vapor diffusion method setup in a 96-well plate. Drops were monitored using a Discovery V8 binocular microscope and an AxioCam ERc5S camera (Zeiss, Germany). Crystals were obtained in the MDL screen, in a condition containing 0.2 M Sodium Acetate, 0.1 M Sodium Cacodylate pH 6.5 and 25% PEG 8000, and using a 2:1 (protein:reservoir) ratio (200 nl:100 nl). In order to obtain bigger crystals, this condition was optimized using the hanging drop method. The final condition (0.2 M Sodium Acetate, 0.1 M Sodium Cacodylate pH 6.5 and 25% PEG 8000), and using a 2:1 (protein:precipitant) ratio (500 nl:250 nl), led to the apparition of three-dimensional crystals (around 75–100 μm). Reproducible crystals appeared after 3 days at 298 K.

### *Data collection and structure resolution of sHPBP*

The crystal was transferred few seconds in a drop (1 μl) containing a cryo-protectant solution made out of the mother liquor plus 10% (v/v) of glycerol. After mounting on a CryoLoop (Hampton research), crystal was flash-frozen in liquid nitrogen. X-ray diffraction intensities were collected on the ID23-1 beamline at the ESRF (Grenoble, France) using a wavelength of 0.97655 Å and a Pilatus 6M detector with 0.15 s exposures. Diffraction data were collected from 1027 images; each frames consisted of 0.15° step oscillations, over a range of 154.05° (Table 1). The molecular replacement using the HPBP structure as model (PDB: 2V3Q) was performed with *Phaser* [38]. The solution was then used for refinement performed using *REFMAC5* and *phenix* [39,40], the model was improved using *Coot* [41]. The model and structure factor were deposited under the Protein Data Bank code PDB: 4M1V (Table 1).

### *HIV inhibition by HPBP and sHPBP*

Inhibition assays were performed in HeLa cells, as previously described [23]. Briefly, HeLa cells were maintained in DMEM + 10 % FBS with antibiotics (100 U/ml of Penicillin and 100 mg/ml of Streptomycin). Cells were pre-incubated with proteins (0.25 μg/ml) for 48 h, then transfected using Lipofectamine 2000 transfection reagent with 0.5 μg of HIV-LTR-luciferase reporter plasmid. Each transfection was done in triplicate.