Journal of Structural Biology 183 (2013) 320–328

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Detection of soluble co-factor dependent protein expression *in vivo*: Application to the 4′-phosphopantetheinyl transferase PptT from *Mycobacterium tuberculosis*

Karine Rottier [a,b,1], Alexandre Faille [a,b,1], Thomas Prudhomme [a,b], Cécile Leblanc [a,b], Christian Chalut [a,b], Stéphanie Cabantous [c,d,e], Christophe Guilhot [a,b], Lionel Mourey [a,b], Jean-Denis Pedelacq [a,b,*]

[a] CNRS, IPBS (Institut de Pharmacologie et de Biologie Structurale), 205 Route de Narbonne, BP 64182, F-31077 Toulouse, France
[b] Université de Toulouse, UPS, IPBS, F-31077 Toulouse, France
[c] INSERM UMR 1037, Cancer Research Center of Toulouse, 20–24 Rue du Pont St. Pierre, 31052 Toulouse Cedex, France
[d] Université de Toulouse, 31052 Toulouse Cedex, France
[e] Institut Claudius Regaud, 31052 Toulouse Cedex, France

ARTICLE INFO

Article history:
Received 13 May 2013
Received in revised form 24 July 2013
Accepted 25 July 2013
Available online 31 July 2013

Keywords:
Split-GFP
Domain trapping
Co-factor
4′-Phosphopantetheinyl transferase
Polyketide
Tuberculosis

ABSTRACT

The need for early-on diagnostic tools to assess the folding and solubility of expressed protein constructs *in vivo* is of great interest when dealing with recalcitrant proteins. In this paper, we took advantage of the picomolar sensitivity of the bipartite GFP1–10/GFP11 system to investigate the solubility of the *Mycobacterium tuberculosis* 4′-phosphopantetheinyl transferase PptT, an enzyme essential for the viability of the tubercle bacillus. *In vivo* and *in vitro* complementation assays clearly showed the improved solubility of the full-length PptT compared to its N- and C-terminally truncated counterparts. However, initial attempts to purify the full-length enzyme overexpressed in *Escherichia coli* cells were hampered by aggregation issues overtime that caused the protein to precipitate within hours. The fact that the naturally occurring Coenzyme A and Mg^{2+}, essentials for PptT to carry out its function, could play a role in stabilizing the enzyme was confirmed using DSF experiments. *In vitro* activity assays were performed using the ACP substrate from the type I polyketide synthase PpsC from *M. tuberculosis*, a 2188 amino-acid enzyme that plays a major role in the virulence and pathogenicity of this microbial pathogen. We selected the most soluble and compact ACP fragment (2042–2188), identified by genetic selection of in-frame fragments from random library experiments, to monitor the transfer of the P-pant moiety from Coenzyme A onto a conserved serine residue of this ACP domain.

1. Introduction

Ways to reduce the toxicity and improve the solubility of recombinant proteins in heterologous hosts have been the subject of prolific research over the past two decades. Major progresses have been made with the advent of structural genomics initiatives and the development of tools to circumvent this major obstacle. Available strains and expression vectors with different promoters now make possible parallel approaches to increase the chance of isolating protein targets in a form suitable for functional and structural studies. Oftentimes, recombinant proteins partition into "inclusion bodies" as a result of misfolding, aggregation, and intracellular accumulation. With the advent of DNA polymerase-based error-prone PCR protocols (Cadwell and Joyce, 1994) as well as *Escherichia coli* mutator strains (Greener et al., 1997), introducing mutations that enhance solubility and stability of a target protein while retaining its activity has now become possible. Wider sequence variations introduced by DNA shuffling techniques (Stemmer, 1994; Zhao et al., 1998) have expanded the functional diversity and facilitated the identification of variants with superior crystallization propensity (Keenan et al., 2005). Libraries of se-

Abbreviations: ACP, acyl carrier protein; AcpS, acyl carrier protein synthase; AnTet, anhydrotetracycline; BSA, bovine serum albumin; CoA, Coenzyme A; DHFR, dihydrofolate reductase; DSF, differential scanning fluorimetry; DTT, dithiothreitol; GFP, green fluorescent protein; IMAC, immobilized metal ion affinity chromatography; IPTG, isopropyl β-D-1-thiogalactopyranoside; Kan, kanamycin; LB, Luria–Bertani; MBP, maltose binding protein; PCR, polymerase chain reaction; PDIM, phtiocerol dimycocerosate; PKS, polyketide synthase; P-pant, 4′-phosphopantetheine; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SEC, size-exclusion chromatography; Sfp, surfactin synthetase-activating enzyme; Spec, spectinomycin; SUMO, small ubiquitin-related modifier.

* Corresponding author at: Université de Toulouse, UPS, IPBS, F-31077 Toulouse, France.
E-mail address: Jean-Denis.Pedelacq@ipbs.fr (J.-D. Pedelacq).
[1] These authors contributed equally to this work.

1047-8477/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsb.2013.07.010

quences fused to reporters with specific marker phenotype (fluorescence in the case of GFP reporters) may be screened to detect soluble variants. Successful results were obtained with the GFP (Pedelacq et al., 2002; Waldo, 2003), although the bulky C-terminally fused reporter may affect the solubility of the passenger protein. Insertion into the GFP scaffolding or fusion to GFP11 prior to complementation with the large fragment GFP1–10 (Cabantous et al., 2005) can overcome this limitation.

When a full-length protein fails to maintain in a soluble form, N- and C-terminal truncations generated by PCR amplification can help identifying variants with increased solubility level (Pedelacq et al., 2011). Another alternative is to select soluble fragments encompassing one or several domains. This can be achieved when domain boundaries are easily predictable from known homologues. The situation becomes more difficult with protein targets that have no sequence or structural homologues. One way to tackle this problem is to generate libraries of truncated DNA fragments coupled with a genetic screen to discriminate between the thousands of constructs. The expression of soluble proteins by random incremental truncation (ESPRIT) (Yumerefendi et al., 2010) and colony filtration (CoFi) blot (Cornvik and Dahlroth, 2006) methods have their limits, since they only permit to truncate unidirectionally one end of the gene. The bidirectional truncation version of ESPRIT (An et al., 2011) and the combinatorial domain hunting (CDH) (Reich et al., 2006) method have no filtering strategy over fragment orientation and frame selection. Upstream protein fusions to murine dihydrofolate reductase (mDHFR) (Dyson et al., 2008) or the β-lactamase (Fisher et al., 2006) can help eliminating incorrect reading frames. However, false positives were identified from translation initiation at internal ribosome binding sites (IRBS). A number of bipartite selection systems have also been developed in an attempt to overcome this issue. The DNA sequence of the target gene is inserted between the two halves of the reporter, which are both required to give an observable phenotype (Cabantous et al., 2008; Daugelat and Jacobs, 1999; Gerth et al., 2004). Our insertional DHFR system combined with the split-GFP assay has been successfully applied to identify soluble fragments from each domain of human p85α and PpsC (Pedelacq et al., 2011).

In this paper, we adopted a PCR approach and a library screening strategy to identify soluble constructs of an enzyme and its protein substrate, respectively. We extend the use of the split-GFP complementation assay (Cabantous et al., 2005) as a diagnostic tool to investigate the solubility of co-factor dependent proteins expressed *in vivo*. We applied this system to the *Mycobacterium tuberculosis* PptT, an enzyme required for growth and persistence of the bacteria *in vivo* (Leblanc et al., 2012). PptT is responsible for the covalent transfer of the P-pant group of Coenzyme A (CoA) to a conserved serine residue onto the acyl carrier protein (ACP) domain of various type I polyketide synthases (PKS) (Quadri et al., 1998a). This post-translational modification converts PKS into functional *holo* enzymes that play a key role in the biosynthesis of various lipids of the mycobacterial cell envelope (Chalut et al., 2006). Split-GFP complementation assays indicated that the full-length enzyme is the most soluble among all the variants tested. A soluble fragment containing the predicted boundaries for the PpsC ACP substrate identified using our domain trapping method (Pedelacq et al., 2011) was used for *in vitro* activity assays. We showed that the natural endogenous CoA and divalent metal ion Mg^{2+}, which play a crucial role for PptT to carry out its function, are also a prerequisite for its stability overtime. This easy-to-use split-GFP system can be generalized to detect soluble co-factor dependent protein expression *in vivo* whenever the presence of co-factors is required for folding and stability.

2. Materials and methods

2.1. PptT small scale expression and solubility tests

The *pptT* gene from *M. tuberculosis* was amplified from genomic DNA using top and bottom primers listed in Supplementary Table 1. The *Nde*I/*Bam*HI digested *pptT* was then inserted into commercially available pET28 vectors (Novagen, Madison, WI, USA) with N- or C-terminal 6His tag. pET vectors allowing the production of PptT with MBP (Leblanc et al., 2012) or SUMO (Invitrogen, Carlsbad, CA, USA) fused to its N-terminus were also tested. Ligated plasmids were transformed into chemically competent *E. coli* BL21(DE3) cells (Invitrogen, Carlsbad, CA, USA). Transformed cells were plated onto Luria–Bertani (LB) agar plates containing 35 μg/ml kanamycin, allowing them to grow overnight at 32 °C. The resulting clones were grown at 37 °C in 1 ml cultures using 35 μg/ml kananycin. Cells were induced in exponential phase with 1 mM IPTG for 3 h. Cell culture pellets of 1 ml of each fragment were separately resuspended in 40 μl TNG buffer (150 mM NaCl, 100 mM Tris–HCl pH = 7.5, 10% (v/v) glycerol) and sonicated. The lysate was fractionated by centrifugation to yield the soluble and pellet fractions. The pellet fraction was washed twice with 100 μl TNG buffer, centrifuged and resuspended in the same starting volume. Samples corresponding to the soluble (S) and pellet (P) fractions were resolved on a 4–20% gradient Criterion SDS–PAGE gel (Bio-Rad, Hercules, CA, USA). Protein samples were stained using Gel Code Blue stain reagent (Pierce, Rockford, IL, USA) and imaged using a GS-800 Calibrated Densitometer (Biorad, Hercules, CA, USA).

2.2. Cloning and in vivo split-GFP solubility screen

Full-length and truncated variants of the *M. tuberculosis PptT* gene were amplified by PCR using a series of forward and reverse primers listed in Supplementary Table 1. Cleaned inserts from *Nde*I/*Spe*I restriction digests were ligated into the pTET ColE1 GFP11 vector and transformed into chemically competent BL21(DE3) cells containing the pET GFP1–10 plasmid (Cabantous and Waldo, 2006; Cabantous et al., 2005). Frozen cells were used to grow 1 mlLuria–Bertani (LB) medium in a 96-well deep-well plate containing 35 μg/ml kanamycin and 112 μg/ml spectinomycin until OD_{600} reached ~1.0 at 37 °C. 50 μl of two successive 400-fold dilutions of the expressed constructs were plated onto a compartmentalized nitrocellulose membrane and grew overnight at 32 °C. The membrane was transferred to an LB/Agar plate containing the same antibiotics supplemented with 250 ng/ml anhydrotetracycline (AnTET) for 2 h at 32 °C, and then moved back onto the original plate for 1 h at the same temperature. Finally, the membrane was transferred onto a plate with same antibiotics and 1 mM IPTG to induce the expression of the GFP1–10 fragment at 32 °C. Colonies were illuminated using an Illumatool Bright light System LT-9900 (http://www.lightools.com/) at excitation and emission wavelengths of 470 nm and 515 nm, respectively. Pictures were taken after 30 min complementation at 32 °C.

2.3. In vitro split-GFP complementation assay

5 ml LB cultures with 112 μg/ml spectinomycin were grown at 37 °C until OD_{600} reached 0.5–0.7. After induction at 32 °C with AnTET at a final concentration of 8 μg/ml and growth for an additional 4 h, cells were harvested by centrifugation at 20,000*g* for 20 min. 20 μl of PptT-GFP11 soluble fractions were mixed with 180 μl of 0.35 mg/ml refolded GFP1–10 in a 96 well microplate (Nunc-Immuno plate, Nunc, Rochester, NY, USA), as previously described (Cabantous and Waldo, 2006). Fluorescence kinetics (λ_{exc} = 488 nm/λ_{em} = 530 nm) were monitored with a