Protein Expression and Purification 141 (2018) 7–18

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Engineering and purification of a thermostable, high-yield, variant of PfCRT, the *Plasmodium falciparum* chloroquine resistance transporter

David J. Wright, Marc O'Reilly, Dominic Tisi*

Astex Pharmaceuticals, 436 Cambridge Science Park, Milton Road, Cambridge CB4 0QA, UK

ARTICLE INFO

Article history:
Received 18 March 2017
Received in revised form
16 June 2017
Accepted 10 August 2017
Available online 17 August 2017

Keywords:
PfCRT
Membrane protein
Membrane transport
Recombinant protein expression
Protein purification
Construct screening
Protein stability
Drug resistance
Malaria
Transporter
Plasmodium falciparum
Neospora canium
NcCRT

ABSTRACT

Historically chloroquine was used to treat the most deadly form of malaria, caused by the parasite *Plasmodium falciparum*. The selective pressure of chloroquine therapy led to the rapid emergence of chloroquine resistant parasites. Resistance has been attributed to the *Plasmodium falciparum* Chloroquine Resistance Transporter (PfCRT), an integral membrane protein of unknown structure. A PfCRT structure would provide new insights into how the protein confers chloroquine resistance and thereby also yield novel opportunities for developing anti-malarial therapies.

Although PfCRT is an attractive target for characterisation and structure determination, very little work has been published on its expression and purification. Here we present a medium throughput protocol, employing Sf9 insect cells, for testing the expression, stability and purification yield of rationally designed PfCRT mutant constructs and constructs of a PfCRT orthologue from *Neospora caninum* (NcCRT). We have identified a conserved cysteine residue in PfCRT that results in elevated protein stability when mutated. Combining this mutation with the insertion of T4-lysozyme into a specific surface loop further augments PfCRT protein yield and thermostability. Screening also identified an NcCRT construct with an elevated purification yield. Furthermore it was possible to purify both PfCRT and NcCRT constructs at milligram-scales, with high purities and with size exclusion chromatography profiles that were consistent with monodispersed, homogeneous protein.

1. Introduction

Malaria is a major threat to worldwide health and was responsible for over 400,000 deaths in 2015 alone [1]. Although numerous anti-malarial treatments have been developed, none of these have been successful in eradicating this disease. Chloroquine is one such treatment and is thought to act by preventing the polymerisation and detoxification of haem degradation products [2]. Chloroquine initially provided a highly effective treatment for malaria; however, parasitic strains less sensitive to chloroquine quickly emerged. It was shown that the resistant strains contained mutations in the gene *pfcrt* [3]. The product of this gene, PfCRT, resides in the membrane of the parasite's digestive vacuole and, when mutated, results in reduced accumulation of chloroquine in this subcellular compartment [4]. Heterologously expressed PfCRT has been shown to transport chloroquine and other antimalarial drugs [5–7], but the native substrate for this transporter remains unknown. PfCRT has broader clinical relevance as malarial strains harbouring a knockout of this gene are non-viable [8], suggesting that even the wild type protein might be a putative drug target. Despite years of study, little is known about the molecular mechanism of PfCRT: in fact, there is still debate as to whether PfCRT acts as a passive or active transporter [9]. Therefore the characterisation and structure of this transporter would be of interest to both academia and the pharmaceutical industry.

Integral membrane proteins are typically challenging targets for structure determination. Although it is estimated that 20–30% of all open reading frames (ORFs) code for membrane proteins [10], less than 2% of structures in the Protein Data Bank (PDB) are of membrane proteins [11]. Many factors contribute to this lack of structural knowledge, but the low stability and purification yield of membrane proteins are perhaps the main reasons. Researchers have adopted two main approaches to enhance the likelihood of being able to solve the three-dimensional structure of a membrane protein. The first is to search for orthologous proteins with

* Corresponding author.
E-mail address: Dominic.tisi@astx.com (D. Tisi).

http://dx.doi.org/10.1016/j.pep.2017.08.005
1046-5928/

increased stability, or purification yield, and the second is to engineer the target membrane protein to have enhanced stability. Indeed, it has been shown that alanine scanning mutagenesis coupled with a radioligand binding, thermostability, assay can be used to engineer a GPCR [12–14], or transporter [15–17], to have increased thermostability and that this correlates with improved crystallisation properties.

An alternative strategy to increase the probability of membrane protein crystallisation is to increase the hydrophilic surface area available for the formation of crystal lattice contacts. This can be achieved by binding a soluble partner protein, such as an antibody [18] or camelid nanobody [19,20], or via a fusion with a soluble partner protein such as T4-lysozyme. There have been particular successes using N-terminal, or intra-cellular loop 3, T4-lysozyme insertions [21–23] or, more recently, modified apocytochrome B_{562} insertions [24–26] with GPCRs. Inserted soluble proteins have been shown to form the majority of crystal contacts in some crystals and this method has been particularly helpful when combined with lipidic cubic phase (LCP) crystallisation (reviewed in Ref. [27]).

In order to identify effectively membrane protein mutations that improve the biophysical properties of a protein it is critical to have a suitably sensitive assay. Although the thermostability and radioligand binding assay approach has been very successful, it is necessary to have access to a high affinity (nM), radiolabelled, ligand for each protein tested. This can be challenging for membrane protein transporters which typically bind ligands, or substrates, with μM-mM affinities [28–30]. Therefore there is a clear requirement for alternative profiling strategies for membrane proteins for which no high affinity ligand is known and for which a structure determination is sought. One method that has been used to estimate the stability of membrane proteins in the absence of a ligand or substrate is fluorescence-detection size exclusion chromatography (FSEC). This technique measures the fluorescence of a green fluorescent protein (GFP) tagged protein in a lysate or complex mixture [31]. A heating step can be added to use this technique to rank protein variants by apparent thermostability [32]. Additionally, there has been significant success using C-terminally appended GFP tags to indicate if a membrane protein has been correctly processed and inserted into the plasma membrane of yeast, or prokaryotic, cells [33]. It also has been shown that GFP is SDS resistant and remains folded when analysed by SDS-PAGE [34].

Although there is much academic interest in PfCRT, there is very little literature precedence for its purification. At the time of publishing, PfCRT had only been heterologously expressed and purified from *Pichia pastoris* [7] or as a double fusion protein in *E. coli* [5]. There are no published protocols for insect cell based purifications and no evidence of construct screening. Given the relatively sparse literature available on PfCRT we also elected to investigate expression and purification of the closest, non-plasmodial, orthologue of PfCRT from *Neospora canium* (NcCRT).

In this work we have employed a medium throughput insect cell expression screen to identify a set of rationally designed PfCRT and NcCRT constructs that have higher purification yield and increased protein thermal stability. We have used a combination of FSEC analysis and measurement of purification yield to further engineer and triage a set of PfCRT and NcCRT constructs for crystallisation screening.

Further we show that some of the mutations and truncations are additive in terms of increasing protein purification yield and stability. We have used FSEC to profile the stability of a PfCRT mutant in a range of detergents and have used these data to identify which detergents might be most suitable for further, large-scale, purification of the PfCRT variants. These data were also used to inform the parameters employed during milligram-scale purifications of high quality PfCRT and NcCRT samples that were subsequently progressed for crystallisation screening and biophysical characterisation.

2. Results

Initial construct design - A full length, wild type, construct for PfCRT (Hb3) was synthesised with a C-terminal thrombin cleavage site, enhanced GFP (eGFP) and deca-histidine tag (Fig. 1 - construct A) and cloned into the pFastBac1 vector. Viruses were produced and PfCRT expression in *Sf*9 insect cells was evaluated by in-gel GFP-fluorescence. Wild type (construct A) protein from 10 ml of *Sf*9 insect cells could be purified using TALON resin and visualised on a Coomassie stained gel (Fig. 1).

Mutation screening- By using the TMHMM [35] transmembrane prediction and RONN disorder prediction [36] servers and an alignment of PfCRT plasmodial orthologues it was possible to design rationally a set of PfCRT constructs to be tested for improved purification yield and thermal stability. The variants selected included deletions, insertions and point mutations (Table 1). These variants were synthesised in the wild type full length PfCRT (Hb3) background, with a C-terminal thrombin cleavage site, enhanced GFP (eGFP) and deca-histidine tag as described previously (Fig. 1 – construct A).

This set of mutants involved modification of regions predicted to be disordered, non-transmembrane, loop regions. Of particular interest were the putative loops between transmembrane helices 2 and 3, or 7 and 8. The only disordered intramembrane region predicted by the RONN disorder prediction server [36] was found between helices 2 and 3 (residues 108–122 inclusive with a >0.5 probability of disorder). Sequence analysis indicated that the region between transmembrane helices 7 and 8 should be ordered and that it contained a cys rich motif that is highly conserved across even distant orthologues (Supplementary Table 1). Thus it was deemed interesting to examine the role that this conserved region plays in maintaining protein stability.

A possible glycosylation site (Asn88), predicted by the NetNGlyc server [37], was also targeted with the mutation N88D. T4 lysozyme

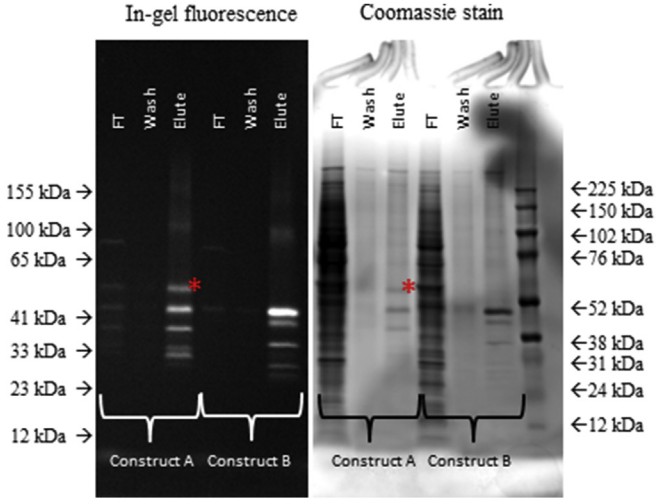

Fig. 1. Purifications of PfCRT truncations. PfCRT constructs were purified from 10 ml of *Sf*9 cells and analysed by SDS-PAGE by in-gel fluorescence (left) and Coomassie staining (right). **Left of each image**: Full length PfCRT followed by a thrombin cleavage site, eGFP and His10 (Fig. 1–construct A) **Right of each image:** Same construct with Δ2-50 and Δ406-424 (construct B). Lanes for each construct correspond to the flow through, 50 mM imidazole wash and elution in 1M imidazole. Marker positions on the left are BenchMark Fluorescent Protein Standard and right are Amersham ECL Full-Range Rainbow Molecular Weight Markers. The positions of "full length" construct are marked with a star (*).