Protein Expression and Purification 129 (2017) 18–24

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Going native: Complete removal of protein purification affinity tags by simple modification of existing tags and proteases

Hui Chin Goh [a], Radoslaw M. Sobota [b], Farid J. Ghadessy [a, *], Saurabh Nirantar [a, **]

[a] p53Lab, Agency for Science Technology and Research (A*STAR), 8A Biomedical Grove, 138648, Singapore
[b] Systems Structural Biology Group, Institute of Molecular and Cellular Biology (A*STAR), 61 Biopolis Dr, 138673, Singapore

ARTICLE INFO

Article history:
Received 6 May 2016
Received in revised form
18 August 2016
Accepted 5 September 2016
Available online 8 September 2016

Keywords:
Protein purification
Affinity tags
Protease
Enforced colocalisation
Native protein

ABSTRACT

Protein purification typically involves expressing a recombinant gene comprising a target protein fused to a suitable affinity tag. After purification, it is often desirable to remove the affinity tag to prevent interference with downstream functions of the target protein. This is mainly accomplished by placing a protease site between the tag and the target protein. Typically, a small oligopeptide 'stub' C-terminal to the cleavage site remains attached to the target protein due to the requirements of sequence-specific proteases. Furthermore, steric hindrance can also limit protease efficiency. Here, we show that respectively fusing the interacting ePDZ-b/ARVCF protein-peptide pair to the target protein and a protease enables efficient processing of a minimised sequence comprising only residues N-terminal to the cleavage site. Interaction of the protein-peptide pair enforces proximity of the protease and its minimised cleavage sequence, enhancing both catalysis of a sub-optimal site and overcoming steric hindrance. This facilitates the high yield purification of fully native target proteins without recourse to specialised purification columns.

1. Introduction

Proteins are the workhorses of living systems, catalysing reactions essential for metabolism, detecting and responding to environmental stimuli, providing structural integrity to the cell structure and detecting pathogens among numerous other functions. Many of these functions are of academic or industrial interest, necessitating the purification of the relevant proteins. Typically, this is accomplished by expressing a recombinant version of the protein of interest fused to an affinity tag that enables column purification using an appropriate affinity matrix. Whilst this process enables routine purification of large amounts of target protein, the presence of the affinity tag fused to the protein can interfere with downstream applications or function [1,2]. The tag may be removed by cleavage of a protease site placed between the target protein and the affinity tag, followed by a step to separate the protein and affinity tag. Nonetheless, most sequence-specific proteases have amino acid requirements at both the N- and C-terminal sides of the cleavage site [3], leaving a small "stub" still attached to the protein of interest. This may be unfavourable for certain downstream applications and in the case of therapeutic proteins, it may raise the possibility of an immune response.

Serial chromatography purifications of cellular lysates can be used to isolate proteins of interest based on their size, charge and pI. However this method is often laborious, non-generic and does not guarantee fully pure protein. Researchers have previously developed methods to fully remove any trace of the affinity tag. A system comprising a column-immobilized mutant subtilisin which is inactive in the absence of fluoride ions, and a target protein fused to subtilisin prodomain at its N terminus has been developed [4]. The prodomain-target protein fusion is loaded onto the mutant subtilisin column, non-specific proteins washed off and cleavage at the junction of the prodomain-protein of interest triggered by the addition of fluoride ions at high concentration. This method requires a specialised column and results in the presence of fluoride ions in the eluate which may have to be removed prior to downstream applications.

Another system developed involves the fusion of the protein of interest to a mutant intein and chitin binding domain at the C terminus [5]. The expressed protein is bound to a chitin matrix and the cellular proteins and debris are washed out. Addition of a strong

* Corresponding author.
** Corresponding author.
E-mail addresses: fghadessy@p53Lab.a-star.edu.sg (F.J. Ghadessy), saurabh.nirantar79@gmail.com (S. Nirantar).

http://dx.doi.org/10.1016/j.pep.2016.09.001

reducing agent triggers the intein, thereby severing the link between the intein and the native target protein, which is eluted. This process also requires a specialised column and results in the presence of a strong reducing agent in the elution buffer.

The TAGzyme method from Qiagen uses aminopeptidases to cleave off N terminal histidine tags, while preventing the cleavage of target protein residues by using "stop points" such as glutamine. These "stop point" containing proteins are processed by Qcyclase which converts the N terminal glutamine into pyroglutamate. This can then be cleaved by an aminopeptidase (pGAPase) to yield native protein. However, the necessity for inserting genetically coded "stop points" for most proteins, and a multi-step process necessary for obtaining native protein leaves substantial room for improvement.

We have therefore developed a purification system to generate native (i.e. completely untagged) proteins of interest (Fig. 1). This requires expression of a fusion protein comprising the ePDZ-b affinity clamp domain adjacent to a sub-optimal sequence-specific protease cleavage site (preferred residues at P1' and P2' are removed) followed immediately by the protein of interest. The cognate protease is expressed with a C-terminal tag comprising the ARVCF peptide sequence that binds ePDZ-b very tightly (K_d < 5 nM) [6]. Typically, the presence of methionine at the P1' position in addition to steric hindrance posed by the folded target protein would render cleavage at this site sub-optimal due to compromised protease affinity. However, fusion to the interacting protein-peptide pair enforces proximity, allowing a much faster reaction rate by increasing the effective concentration of protease in the vicinity of a sub-optimal cleavage site. The concept of enforced proximity has previously been described to generate proteins with native amino termini from MBP fusion proteins, although these still retained C-terminal histidine tags [7].

Here, we show that this method leads to increased amounts of native protein and is readily transposable onto pre-existing protocols and reagents used for purification of proteins labelled with the ubiquitous histidine tag.

2. Materials and methods

Unless otherwise specified, all oligonucleotides used in this work were from Integrated DNA Technologies (Singapore), chemical reagents from Sigma and SplB protease from Thermo Scientific. His-tag purification columns were purchased from GE Healthcare.

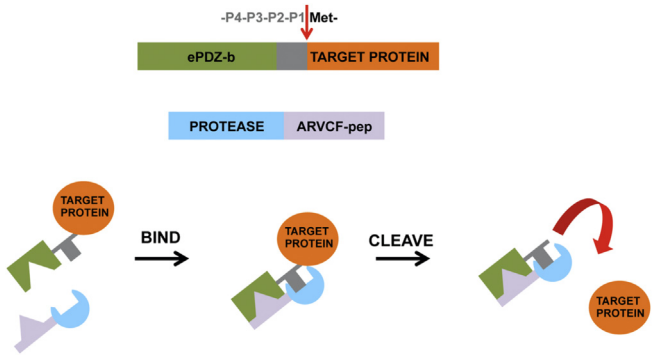

Fig. 1. Optimising cleavage at suboptimal protease sites by enforced co-localisation. Top schematic depicts domain structure of fusion protein comprising target of interest and ePDZ-b connected by protease cleavage site with methionine at P1' position. Protease is expressed with C-terminal ARVCF-pep tag. Interaction of ARVCF-pep (light magenta) with ePDZ-b (green) promotes co-localisation of protease (cyan) with sub-optimal cleavage site (grey), resulting in cleavage and generation of target protein with no exogenous amino acids at the N-terminus. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Endogenous TEV protease comprising N-terminal histidine tag and the MPB-ENLYFQS-PH-G1VCA protein were gifts from the Robinson laboratory.

2.1. Plasmid construction

Expression cassettes were made by PCR amplification of individual genes with oligonucleotides incorporating linker extensions which could then be joined to other genes via Splicing by Overlap Extension (SOE). Fusion genes were inserted into expression vector pET22 (Kan) (pET22b with ampicillin resistance cassette replaced with kanamycin resistance) by infusion cloning. Oligonucleotide sequences are listed in Supporting Information (Table S1).

2.2. Protease gene constructs

SplB proteases were expressed with a protein A domain fused at the N terminus, an SplB cleavage site "WELQ" immediately preceding the first amino acid of the mature SplB protease, and a 6X histidine tag followed by the ARVCF peptide at the C terminus (protein A-GGWELQ-mature SplB-His tag-ARVCF peptide). The protein A is removed by cleavage at the "WELQ" site for a small minority of fusion proteins during protein expression, forming active SplB protease, which in turn cleaves more of the fusion protein, resulting in a rapid chain reaction yielding mature SplB-His tag-ARVCF peptide.

The Protein A open reading frame (ORF) was amplified using oligonucleotides oSN684 and oSN1096. oSN1096 appends a sequence coded as "GGSGKGGWEL" at the 3' of protein A. SplB was PCR-amplified with oSN1099 and oSN875. oSN875 appends a partial His tag to the 3' of SplB. oSN1099 appends a sequence complementary to that of oSN1096, enabling SOE PCR of these two amplicons to yield protein A-WELQ-SplB-partial His tag cassette.

Simultaneously, the pET22b (kan) vector was amplified by inverse PCR using oSN1183 and oSN1184, which have complementary 5' ends for infusion cloning, followed by intra-plasmid infusion cloning (Clontech), creating a vector which has the "PQPVDSWV" coding ARVCF peptide just downstream of the XhoI site. This vector was then linearized with NdeI/XhoI, and the fusion ORF protein A-WELQ-SplB-partial His tag was inserted by infusion cloning yielding protein A-WELQ-SplB-His tag-ARVCF peptide in pET22b (Kan) vector.

The TEV-AP4 protein expression construct encoded the ORF for a maltose binding protein (MBP)-optimal TEV cleavage site (ENLYFQS)-TEV-GGGHHHHHHGG**DSWV** fusion protein. The ORF was synthesized (Genscript) and inserted into the NdeI/XhoI sites in pET22b vector. Note the TEV expressed incorporated the S219V mutation to reduce autolysis [8]. The TEV protease auto-cleaves itself from MBP after expression and is purified using the 6-histidine tag preceding the truncated ARVCF peptide (DSWV, indicated in bold above). The vector pRK793 [8] was used to express histidine-tagged TEV comprising the S219V mutation.

2.3. SplB-His tag-truncated ARVCF peptide constructs

These were constructed by doing inverse PCR with the above-mentioned protein A-WELQ-SplB-His tag-ARVCF construct using the oligonucleotides oSN1225, 1226, 1227, 1228 and 1230 with the common reverse oligonucleotide oSN1229 to yield constructs expressing ARVCF peptides PQPVDSWV, QPVDSWV, PVDWSV, VDSWV and DSWV respectively.

2.4. SplB-truncated His tag-DSWV constructs

The protein A-WELQ-SplB-truncated His tag-DSWV was created