Protein Expression and Purification 149 (2018) 31–36

Contents lists available at ScienceDirect

Protein Expression and Purification

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

A versatile expression vector for the growth and amplification of unmodified phage display polypeptides

Alexander J. Winton, Janae L. Baptiste, Mark A. Allen*

Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, MD 21250, USA

ARTICLE INFO

Keywords:
Phage display
Polypeptide
Expression
Cloning vector
HSQC
GEPI

ABSTRACT

Proteins and polypeptides represent nature's most complex and versatile polymer. They provide complicated shapes, diverse chemical functionalities, and tightly regulated and controlled sizes. Several disease states are related to the misfolding or overproduction of polypeptides and yet polypeptides are present in several therapeutic molecules. In addition to biological roles; short chain polypeptides have been shown to interact with and drive the bio-inspired synthesis or modification of inorganic materials. This paper outlines the development of a versatile cloning vector which allows for the expression of a short polypeptide by controlling the incorporation of a desired DNA coding insert. As a demonstration of the efficacy of the expression system, a solid binding polypeptide identified from M13 phage display was expressed and purified. The solid binding polypeptide was expressed as a soluble 6xHis-SUMO tagged construct. Expression was performed in *E. coli* using auto-induction followed by Ni-NTA affinity chromatography and ULP1 protease cleavage. Methodology demonstrates the production of greater than 8 mg of purified polypeptide per liter of *E. coli* culture. Isotopic labeling of the peptide is also demonstrated. The versatility of the designed cloning vector, use of the 6xHis-SUMO solubility partner, bacterial expression in auto-inducing media and the purification methodology make this expressionun vector a readily scalable and user-friendly system for the creation of desired peptide domains.

1. Introduction

Short proteins and polypeptides constitute an important avenue of research. Various disease states such as the formation of amyloid plaque and prion diseases are a result of the misfolding of short proteins [1,2]. Currently, there are many examples of therapeutics composed of polypeptides, these included antimicrobial peptides and cell targeting moieties [3–7]. In addition to therapeutic, diagnostic, or disease relevance, there also exist biomineralizing and material binding polypeptides; these have been shown to mimic natural biomineralizing systems and have been utilized to improve the properties of inorganic materials by means of tethering [8], material synthesis and engineering [9–13]. Many of these peptides are either synthesized using solid support peptide synthesis or they are isolated (typically in low yield) from the organisms in which they were identified [14–16] and there are many examples of polypeptides being expressed in bacterial expression systems [17–19]. This article focuses on the development of a heterologous expression system that is broadly applicable to preparing nearly any short sequence of amino acids. As an example of utility a specific binding epitope identified through phage display is expressed and verified to yield nearly 10 mg of purified polypeptide per liter of *E. coli*.

There are several examples of polypeptides that have been heterologously expressed in *E. coli* [17–19]. Since short polypeptides are prone to protease degradation inside bacterial cells [20] they must be expressed with a fusion or solubility partner. Solubility partners have been extensively reviewed for their role in expressing small proteins and polypeptides [21–23]. Typical fusion partners include maltose binding protein, thioredoxin, glutathione-S-transferase, and small ubiquitin like modifier (SUMO) proteins [21]. Fusion partners increase the solubility of a desired peptide and also serve as protection from protease cleavage while inside *E. coli* [21,24]. With a fusion partner, polypeptides can be expressed at a high yield, they remain soluble in *E. coli* cells, and are easily isolated through a hexa-histidine affinity tag present on the fusion partner. After expression and isolation of the fusion partner and peptide of interest, the fusion partner must be cleaved away from the peptide using some defined protease [25]. Fusion partners and expression methods play important roles in increasing yields but they often require that the peptide of interest be modified with unwanted amino acids in order to make the expression system applicable. Unwanted amino acid residues have to be inserted for two different reasons; one has to do with the development of an expression vector i.e. which type of restriction cut site is engineered into the vector

* Corresponding author.
E-mail address: allenm@umbc.edu (M.A. Allen).

https://doi.org/10.1016/j.pep.2018.04.008
Received 23 January 2018; Received in revised form 12 April 2018; Accepted 13 April 2018
Available online 17 April 2018
1046-5928/

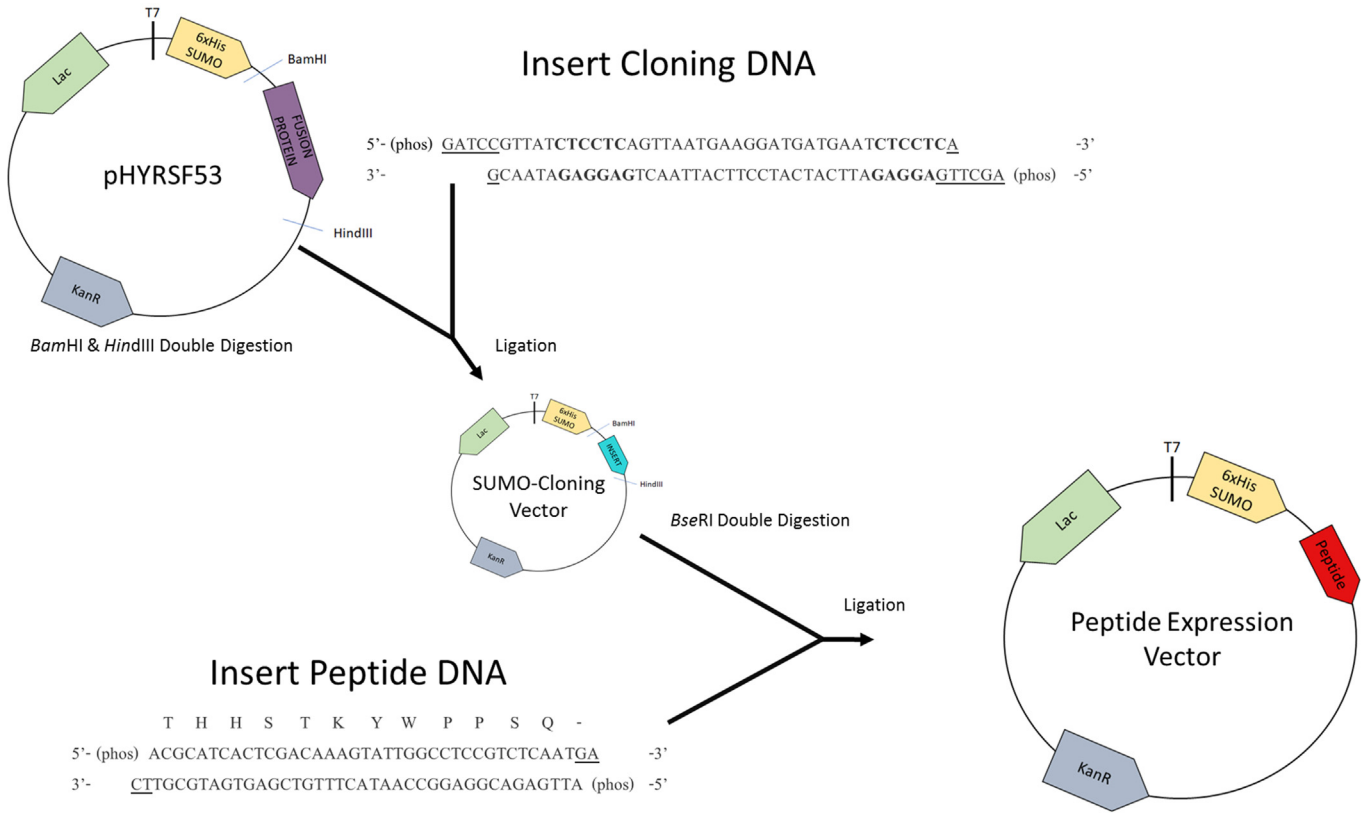

Fig. 1. Diagram of the SUMO-Cloning vector and Peptide Expression vector. The Insert Cloning Sequence containing two *Bse*RI digestion sites (bold) was inserted into the pHYRSF53 vector using *Bam*HI and *Hin*dIII sites. Double digestion of the SUMO-Cloning vector with *Bse*RI followed by insertion of the Insert Peptide DNA results in the Peptide Expression vector. This vector enables the expression of 6xHis-SUMO immediately followed by the peptide of interest.

and the second is related to the type of protease cleavage site that is necessary to remove the polypeptide of interest from the fusion partner and affinity tag. While seemingly unimportant when observing larger proteins, additional amino acid residues may detrimentally impact the performance of a 12 amino acid polypeptide. Many researchers choose to avoid bacterial expression systems by using solid support peptide synthesis. Chemical synthesis enables the creation of unmodified polypeptides, free of additional amino acids. However, solid support synthesis of polypeptides exhibits relatively limited yields and if being used for growing isotopically labeled peptides is an expensive endeavor both in terms of materials and time [26,27]. Heterologous expression offers unique advantages over synthesis. One advantage is that polypeptides can be grown in media that enrich the expressed protein in isotopically labeled amino acids. These isotopically labeled amino acids serve as NMR active handles or as activity reporters in isotope dilution mass spectrometry [26–28]. In addition, point mutations are easily made in the cloning vector which can help elucidate a peptide domain's mechanism of action or to add specific amino acid functionalities without having to redundantly synthesize each different peptide variation. One common technique that is used to identify a polypeptides mechanism of action involves modifying each amino acid in a peptide to some non-functional amino acid, typically alanine, and then observing how the properties of the peptide change. Alanine scanning is commonly performed on functional polypeptides and relies on converting amino acids one by one to alanine, ultimately allowing for the identification of pertinent residues responsible for function [29–31].

The focus of this article is to communicate the generation of a versatile expression system which is broadly applicable to the production of short polypeptides. As a demonstration of the utility of the expression system, a 12mer polypeptide exhibiting a specific binding epitope for lithium titanate oxide (a lithium ion battery anode material) which was identified by phage display is expressed. Among the advantages to using this Peptide Expression vector are; the ability to grow large quantities of polypeptides, the ability to easily modify the polypeptides through mutagenesis, and the ability to easily isotopically label the polypeptide for analysis through NMR spectroscopy or mass spectrometry. To our knowledge, this serves as the first time that solid binding polypeptides identified through phage display were heterologously expressed and cleaved from a fusion partner to provide pure and unmodified polypeptides.

2. Materials and methods

2.1. Materials and reagents

The SUMO plasmid, pHYRSF53, and the ULP1 protease plasmid, pFGET19_ULP1, (were gifts from Hideo Iwai (Addgene plasmid #64696 and #64697 respectively)) and were received from addgene [22]. Chemically Competent BL21 (DE3) *E. coli* cells, SOC outgrowth media, Ph.D.™ -12 Phage Display peptide library kit, *Bam*HI-HF®, *Hin*dIII-HF®, and *Bse*RI restriction enzymes, 10x Cutsmart® Buffer, Gel loading dye Purple 6X, Quick-Load® 2-Log DNA Ladder, Quick Dephosphorylation kit (CIP), and Quick Ligation™ kit were all purchased from New England BioLabs. All DNA primers and inserts were purchased from Integrated DNA Technologies (IDT®). Precision Plus Protein™ Dual Xtra Standards protein ladder was purchased from BioRad. Synthesized THHSTKYWPPSQ polypeptide was purchased from Biomatik (www.biomatik.com) and was received with 92.8% purity. All other reagents were purchased from VWR Life Sciences.

2.2. Creation of SUMO-Cloning vector

DH5α *E. coli* cells containing the vector pHYRSF53 were streaked onto a Lysogeny Broth (LB) Agar plate supplemented with kanamycin.