Protein Expression and Purification 62 (2008) 128–137

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Addressing *Shewanella oneidensis* "cytochromome": The first step towards high-throughput expression of cytochromes *c*

Yuri Y. Londer *, Sarah E. Giuliani, Terese Peppler, Frank R. Collart

Biosciences Division, Argonne National Laboratory, 9700 South Cass Avenue, Building 202, Argonne, IL 60439, USA

ARTICLE INFO

Article history:
Received 16 June 2008
Available online 8 July 2008

Keywords:
Cytochrome *c*
High-throughput
Ligation-independent cloning
Periplasmic expression
Protein expression
Shewanella oneidensis

ABSTRACT

Integrated studies that address proteins structure and function in the new era of systems biology and genomics often require the application of high-throughput approaches for parallel production of many different purified proteins from the same organism. Cytochromes *c*—electron transfer proteins carrying one or more hemes covalently bound to the polypeptide chain—are essential in most organisms. However, they are one of the most recalcitrant classes of proteins with respect to heterologous expression because post-translational incorporation of hemes is required for proper folding and stability. We have addressed this challenge by designing two families of vectors (total of 6 vectors) suitable for ligation-independent cloning and developing a pipeline for expression and solubility analysis of cytochromes *c*. This system has been validated by expression analysis of thirty genes from *Shewanella oneidensis* coding for cytochromes *c* or cytochromes *c*-type domains predicted to have 1–4 hemes. Out of 30 targets, 26 (87%) were obtained in soluble form in one or more vectors. This work establishes a methodology for high-throughput expression of this class of proteins and provides a clone resource for the microbiological and functional genomics research communities.

The capability to express proteins in heterologous systems is an important enabling feature for structural and functional studies of proteins. One of major challenges of post-genomic biology is the development of cost-efficient high-throughput methods for heterologous production of proteins from newly sequenced genomes. Cytochromes *c*, where heme is covalently bound to the polypeptide chain[1], pose an additional challenge for heterologous expression since the nascent apoprotein must undergo correct post-translational modification (covalent heme attachment) to enable functionality and proper folding, which presents a number of challenges. First of all, cytochromes *c* are extracytoplasmic proteins (either they are soluble periplasmic proteins or they are anchored in the inner membrane, facing the periplasm, or they are anchored in the outer membrane). Consequently, both apoprotein and heme must be translocated through the inner membrane. During this process, the cysteines that will covalently bind the hemes must be kept reduced in the oxidative periplasmic environment, while hydrophobic heme must stay in solution. Finally, there must be an enzyme catalyzing formation of the thioether bonds between heme vinyl groups and the cysteines in the characteristic heme-binding motif Cys-Xaa-Xaa-Cys-His [1–4]. All organisms studied to date use one of three pathways, or systems, to address these challenges [3,5]. *Escherichia coli* (as well as many other Gram-negative bacteria) employs System I, which includes an ABC transporter, a thioredoxin-like redox system, a heme chaperone, and a heme lyase. *Escherichia coli* has eight proteins, consecutively named CcmA through CcmH, that are responsible for maturation of cytochromes *c* [1–4]. These proteins are located in the inner membrane and on the periplasmic side of inner membrane. They are encoded by gene cluster *ccmABCDEFGH* that is a part of so-called "*aeg46.5*" operon active only under anaerobic conditions [6]. Cloning of this cluster into pACYC184 (a plasmid compatible with many commercially available expression vectors) allowed its constitutive co-expression under control of the tetracycline promoter [7] and made possible the production of properly matured *c*-type cytochromes in *E. coli* under aerobic conditions ([8,9] and references therein). The only other major requirement for successful synthesis of recombinant cytochromes (besides co-expression of the *ccm* genes) is the presence of a leader peptide to target the nascent polypeptide to the periplasm [10].

Shewanella oneidensis MR-1 is an example of organism with a large number of predicted cytochrome *c* genes [11,12], which may

* Corresponding author. Current address: New England Biolabs, 240 County Road, Ipswich, MA, USA. Fax: +1 630 252 3387.
E-mail address: londer@neb.com (Y.Y. Londer).

[1] Strictly speaking, cytochromes *c* are electron transfer proteins where heme is covalently bound to the polypeptide. This definition excludes other proteins with functions other than electron transfer that contain *c*-type hemes (covalently bound hemes), such as nitrite reductase or a recently described family of sensor proteins [39,41]. However, the *E. coli* expression system described here has been shown to provide covalent heme attachment in many different types of heme *c*-containing proteins, regardless of their function. For the purpose of brevity, we use the term "cytochrome *c*" throughout this manuscript but all statements and conclusions made here should be applicable to all proteins containing *c*-type hemes.

1046-5928/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2008.06.014

…cattggtaactgtcagaccaagtttactcat… → …cattggtaactgtcaga*ccatgg*ttactcat…

…H W * …H W *

Fig. 1. Mutagenesis of pLBM4. Fragments of DNA sequence corresponding to the β-lactamase gene are underlined. *Nco*I restriction site is shown in italic.

be responsible for its respiratory versatility. The ability of this organism to couple its metabolism to the respiration of many different electron acceptors, including soluble and insoluble metal oxide reduction and other anaerobic respiratory pathways [13–16], makes *S. oneidensis* attractive as a model organism to assess bioremediation potential. The ability to reduce and precipitate from solution a variety of different toxic metal ions (e.g., chromium or uranium) suggests that, if the processes were understood in greater detail, metal ion-reducing microbes could be exploited for the bioremediation of metal-contaminated sites. Therefore, studies of individual redox components of metal reducers, such as *S. oneidensis*, are a high priority.

This work outlines a systematic and comprehensive approach for expression of cytochromes *c* using the set of *S. oneidensis* cytochromes predicted to have four or fewer covalently bound hemes (total of 30).The application of a high-throughput strategy for expression of a "cytochromome" (the total of cytochromes *c* encoded in a given genome) is a unique approach that will increase our understanding of the functional capabilities of these proteins. This approach also enabled the design and evaluation of 6 different vectors (grouped in two families) for their potential to be used in high-throughput cytochrome *c* production.

Materials and methods

Bacterial strains and plasmids

Escherichia coli strains XL1-Blue (Stratagene) and DH5α (Invitrogene) were used for subcloning. BL21(DE3) (Novagen) co-transformed with plasmid pEC86 [7] was used for protein expression. Growth media were supplemented with carbenicillin, 100 μg/ml, and chloramphenicol, 34 μg/ml, where appropriate. Plasmid pEC86 containing the *ccm* gene cluster was a kind gift from Dr. L. Thöny-Meyer (ETH, Zürich, Switzerland).

Vector design

DNA manipulations generally followed standard published procedures [17]. Oligonucleotides for mutagenesis and vector design were synthesized by MWG Biotech (High Point, NC) and are listed in Table 1. KOD DNA polymerase (EMD Biosciences) was used for PCR amplifications. PCR products were purified using QIAquick PCR purification kit or QIAquick Gel extraction kit (Qiagen). Restriction enzymes used were purchased from New England Biolabs. For mutagenesis, QuikChange Site-Directed Mutagenesis kit (Stratagene) was used in accordance with the manufacturer's instructions.

Table 1
Primers used for mutagenesis and vector design

Primer	Sequence (5′-3′)[a]
NcoA-F	gcattggtaactgtcagaccatggttactcatatatactttagattg
NcoA-R	caatctaaagtatatatgagtaaccatggtctgacagttaccaatgc
SkpA-F	actacatgtaaggagtttattgtgaaaaagtgg
SkpA-R	accacatgtacttatttaacctgtttcagtac
FkpA	cataaacatgtggagatatggatgaaatcactg
FkpR	atttaacatgttatttttttagcagaatctgcggc
HisF	gatggcccaccatcatcatcatcattc
T7-Ter	gctagttattgctcagcgg
T7-Pro	taatacgactcactataggg
PelR	gatggtgggccatcgccggctg
MBP	cgatggccaaaatcgaagaaggtaaactgg
MBP-R	ccagtttaccttcttcgattttggccatcg
K-Up	cacaggaaacagctatgacc
C7-R	ttataggtaccagaagaacccttcttgtggcactcgc
Amp-F	cggcaacaattaatagactgg

[a] Underlined are restriction sites (where appropriate) and the overlapping sequence in primers HisF and PelR.

The first group of vectors derived from pLBM4 [9]. To make pLSM2 and pFCM21, site-directed mutagenesis with primers NcoA-F and NcoA-R was used to introduce a unique NcoI restriction site immediately following β-lactamase gene (Fig. 1). A DNA fragment including the *E. coli skp* gene [18–20] and its native ribosome binding site was amplified from the genomic DNA with primers SkpA-F and SkpA-R introducing AflIII restriction sites on both sides. The resulting DNA fragment was digested with AflIII and cloned into the mutated pLBM4 digested with NcoI. After ligation, the DNA-ligase was inactivated by incubation at 80 °C for 20 min and the mixture was treated with NcoI once again to get rid of recircularized vector molecules. AflIII and NcoI form compatible cohesive ends but a hybrid site obtained after their ligation cannot be recognized by either of the enzymes. However, in recircularized vectors the NcoI site is restored and treatment with this enzyme should turn them into linear molecules that transform *E. coli* cells very inefficiently. Clones obtained after transformation were tested by PCR to select for those containing the insert in a proper orientation. PCR was performed with primers SkpA-R and Amp-F. Amp-F anneals to the β-lactamase gene ∼300 bp upstream of the NcoI site so that in the case of properly positioned insert a PCR product of 793 bp was generated, while no PCR product in the cases of empty vector and/or inversely oriented insert. The resulting construct provided constitutive expression of Skp under control of the β-lactamase promoter from polycistronic mRNA.

To make pFCM21, a DNA fragment containing the gene of *E. coli* peptidylprolyl isomerase FkpA [21,22] and its native ribosome binding site was amplified from the genomic DNA with primers FkpA and FkpR and cloned into the NcoI site of mutated pLBM4 following the same approach.

The second group of vectors was derived from pMCSG9 [23]. A fragment of pMCSG9 containing His-tag, MBP[2], TEV cleavage site, LIC site and 3′-untranslated region including HindIII restriction site was amplified with primers HisF and T7-Ter. Then, a fragment of vector pET-22b(+) (Novagen) encoding the PelB leader sequence and adjacent 5′-untranslated region including XbaI restriction site was amplified with primers T7-Pro and PelR. Primers HisF and PelR contained complementary sequences of 14 bases (underlined). Therefore, the two PCR fragments overlapped and were recombined by overlap PCR. The resulting PCR fragment was digested with restriction enzymes XbaI and HindIII and ligated into pASK40 [24] digested with the same enzymes. Then two rounds of site-directed mutagenesis were carried out to remove the two SspI sites in the body of plasmid. The resulting vector was named pMKL1.

To make pMTL4, a fragment of pMKL1 that codes for the His-tag and linker between the tag and the mature MBP sequence was removed by deletional mutagenesis with primers MBP and MBP-R.

To make pCSF23, a fragment of pCK32 [25] that contains 5′-untranslated region (including XbaI restriction site), OmpA leader sequence and mature cytochrome c_7 sequence was amplified with primers K-Up and C7-R. The latter primer introduced restriction site KpnI (underlined) and a short linker instead of the stop codon.

[2] *Abbreviations used:* LIC, ligation-independent cloning; MBP, maltose-binding protein; TEV protease, tobacco etch virus protease.