Analytical Biochemistry 512 (2016) 110–113

Contents lists available at ScienceDirect

Analytical Biochemistry

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

Technical note

Construction of a compatible Gateway-based co-expression vector set for expressing multiprotein complexes in *E. coli*

Loubna Salim [1], Claire Feger [2], Didier Busso*

Service de Biologie Moléculaire, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/Inserm/Université de Strasbourg, 1 rue Laurent Fries, BP10142, 67404 Illkirch cedex, France

ARTICLE INFO

Article history:
Received 27 June 2016
Received in revised form
17 August 2016
Accepted 20 August 2016
Available online 21 August 2016

Keywords:
Co-expression
Expression screening
Gateway
Protein complex
Recombinational cloning
Versatility

ABSTRACT

We report the construction of a versatile Gateway-based co-expression vector set for producing multiprotein complexes in *Escherichia coli*. The set consists of two groups of three vectors (pCoGW and pCo0GW), each having a specific antibiotic resistance gene, a compatible origin of replication and allowing cloning of up to two genes, each under control of its own T7 promoter. To validate the set, 33 (co-)expression plasmids encoding fluorescent protein (GFP, DsRed and ECFP) have been generated. Protein expression levels were quantified and (co-)expression visualized by fluorescent microscopy. The results illustrate the applicability of these vectors in co-expression studies.

Multiprotein complexes catalyze vital processes in the cell. Although isolation and structure determination of endogenous complexes has been reported [1], *in vitro* reconstitution by co-expressing proteins appears to be the appropriate compromise for overproduction of soluble and functional multiprotein complexes at high yields for further biological and/or structural studies.

Many different prokaryotic and eukaryotic heterologous expression systems have been reported (see Methods in Molecular Biology, volume 824), however, the bacterium *Escherichia coli* is still used as the principal expression system since lot of tools, improvements and developments have been made for producing recombinant proteins [2]. Moreover, different strategies and several vector sets have been recently developed for protein complexes production in *E. coli* [3].

Co-expression can be conducted following different strategies. One strategy consists of co-transforming compatible vectors, each one encoding for a single protein [4–9]. The limitation of this strategy is that the number of compatible origin of replication and the number of distinct antibiotic resistance genes is limited, resulting in a reduced number of proteins being able to be expressed simultaneously (*i.e.* not more than 4). A second strategy consists of subcloning genes in the same plasmid, such that each gene is being controlled by a separate promoter, or, alternatively, a unique promoter is responsible for the expression of an operon like transcript [6,8,10]. Limitations of such a strategy are linked to the size of the plasmid and to the restriction-ligation cloning technology often used to subclone genes. Indeed, the frequency to meet the restriction sites for subcloning additional gene increases with the number of already cloned genes. To circumvent this problem, gene concatenation from an initial plasmid might be accomplished using rare-cutter restriction sites [6] or by inserting additional genes by restriction-free or PCR-based strategies [11]. To avoid handling too large plasmid, both strategies could be combined to increase the number of proteins to be expressed simultaneously in the cell.

The aim of this study was to develop a set of co-expression vectors that would circumvent some of the issues described above. To this end, the vectors would have to pass a certain number of criteria: (i) restriction of the plasmid size by cloning no more than two genes per plasmid; (ii) allow each gene to be under

Abbreviations: ECFP, enhanced cyan fluorescent protein; DNA, deoxyribonucleic acid; GFP, green fluorescent protein; IPTG, isopropyl-beta-thio-galactoside; MCS, multiple cloning site; PCR, polymerase chain reaction.

* Corresponding author. Present address: CEA/DRF/IRCM/CIGEx, 18 route du Panorama BP6, 92265 Fontenay-Aux-Roses, France.
E-mail address: didier.busso@cea.fr (D. Busso).
[1] Present address: Lycée franco-qatari Voltaire, Po Box 12634, Doha, Qatar.
[2] Present address: Laboratoire de Diagnostic génétique NHC, 1 place de l'hôpital, BP 426, 67091 Strasbourg Cedex, France.

http://dx.doi.org/10.1016/j.ab.2016.08.018
0003-2697/

Table 1
List of elaborated pCoGW and pCo0GW vectors.

Vector name	N-ter His6	Resistance	Ori	Based on	GeneBank#
pCoGWA	Yes	Ampicillin	ColE1	pETDuet-1	JX262231
pCoGWC	Yes	Chloramphenicol	P15A	pACYCDuet-1	JX262232
pCoGWS	Yes	Spectinomycin	CloDF13	pCDFDuet-1	JX262233
pCo0GWA	No	Ampicillin	ColE1	pETDuet-1	JX262234
pCo0GWC	No	Chloramphenicol	P15A	pACYCDuet-1	JX262235
pCo0GWS	No	Spectinomycin	CloDF13	pCDFDuet-1	JX262236

The pCoGW vectors allow addition of a 6× histidine tag at the N-terminus of the protein encoded by the gene cloned by Gateway under control of the T7 promoter 1.

control of its own T7 promoter for rapid parallel expression screening using auto-inducible strategy [12], (iii) introduction of a N-terminal 6× histidine tag to at least one of the proteins for affinity purification; (iv) creation of a protease site after the 6× histidine tag allowing its cleavage for further functional or structural studies; (v) combination of compatible plasmids having similar copy number; (vi) robust and straight-forward cloning procedure to clone the gene of interest, irrespective of the vector chosen; and (vii) having the possibility to rapidly screen for protein partner candidates for a given target protein. To elaborate the set, we selected the three following Duet vectors (Merck-Novagen - EMD Millipore, Darmstadt, Germany): the pETDuet-1 (ColE1 origin, ampicillin resistance, 40 copies), the pACYCDuet-1 (P15A origin, chloramphenicol resistance, 10–12 copies) and the pCDFDuet-1 (CloDF13 origin, spectinomycin resistance, 20–40 copies). These vectors fit three criteria: (i), (ii) and (v). To fit the criteria (vi) and (vii), we decided to introduce a Gateway conversion cassette (Invitrogen, Carlsbad, CA) at the MCS1 allowing restriction-free cloning by recombination, and to replace the MCS2 by a convenient cloning cassette with the *Nde*I and *Bgl*II restriction sites. Finally, to fit the last criteria (iii and iv) a N-terminal 6× histidine tag encoding sequence has been inserted upstream of MCS1 (pCoGW) that can be removed as well as the translated *att*B1 sequence by designing the desired protease cleavage site encoding sequence in the PCR primer [13].

The vectors have been engineered as described elsewhere (see the Experimental Design, Materials and Methods section and the Figure 1 in Ref. [14]). First, the MCS1 of the selected Duet plasmids has been modified to encode the 6× histidine tag (pCo) or no tag (pCo0). The MCS2 was replaced with a sequence adding *Nde*I, *Bgl*II and few additional cloning sites as well as a sequence encoding a C-terminal Flag tag. Finally, the Gateway conversion cassette (Invitrogen) was inserted resulting in the vector sets that consisted of the pCoGW and the pCo0GW vector series (Table 1). The plasmids were sequence verified. All generated plasmid sequences have been deposited on GeneBank under accession number JX262231 to JX262236.

To validate the vector set, genes encoding GFP (from pcDNA-DEST53 plasmid – Invitrogen), DsRed (from pDsRed2-N1 plasmid – Clontech, Mountain View, CA) and ECFP (from pECFP-C1 plasmid – Clontech) proteins were cloned in pDONOR207 entry plasmid by Gateway [14]. From these plasmids, 30 pCoGW and pCo0GW expression vectors have been constructed including both single gene and double gene constructs (see Table 2 in Ref. [14]). As a control for protein expression, each of the three genes were cloned in the pHGWA plasmid [15]. The vectors were validated according to two principles: (i) protein expression levels should be similar between different vectors independent from vector and gene location (*i.e.* under control of either T7 promoter 1 or T7 promoter 2), and (ii) up to three compatible plasmids should be maintained in the cell without any loss in protein expression levels as compared to that obtained for the single expression experiments.

To conduct the validation experiments, protein expression was imaged and quantified (see the Experimental Design, Materials and Methods section in Ref. [14]). Briefly, expression vector(s) transformed BL21(DE3) strain and transformation mix was split onto two selective agar plates. After an overnight culture, colonies obtained on one plate were replicated on a nitrocellulose filter placed on a selective plate containing IPTG for inducing gene expression. The plate was incubated for an additional day at room temperature and fluorescent protein expression was imaged using the MacroFluo macroscope (Leica, Solms, Germany). In parallel, four colonies from the second plate were picked to inoculate individual starter cultures. The day after, cultures in ZYM-5052 auto-inducible medium [12] were inoculated and grown for 24 h at 25 °C. Upon cultivation, protein expression was quantified by measuring emission fluorescence level using the Mithras multimode microplate reader (Berthold technologies, Bad Wildbad, Germany).

To evaluate the single- and double gene constructs for protein expression, a total of 26 expression experiments were performed in triplicate (see Table 3 in Ref. [14]). Expression of the fluorescent proteins was quantified by measuring emission fluorescence level on the Mithras multimode microplate reader (Berthold technologies, Bad Wildbad, Germany) (see the Experimental Design Materials and Methods section in Ref. [14] for experimental details). GFP and DsRed proteins could be clearly visualized by fluorescence microscopy after co-expression using the pCo0GWA vector as backbone (Fig. 1A). The expression level of each protein proved to be unaffected by the position of the gene within the construct, *i.e.* under control of either T7 promoter 1 or 2 (Fig. 1A and B). In addition, the absolute quantities were similar to the amount of protein expressed from the control vector (pHGWA) (Fig. 1B). Protein expression was also visualized and quantified for the remaining constructs (Figure 1A and 1B and Table 3 in Ref. [14]). Comparison of the single gene constructs and double gene constructs showed that they were equally competent in production of the proteins, demonstrating that adding a second gene to the construct does not compromise recombinant protein expression levels. Furthermore, these data show that there is no discrepancy between the pCoGW and pCo0GW set of vectors when analyzing fluorescent protein expression. This statement was also confirmed by performing an additional set of 26 experiments where each of the pCoGW vectors was co-transformed independently with the two pCo0GW vector allowing protein expression from two genes cloned individually in compatible vectors (data not shown). As a conclusion for this first experiment sets, (i) the position of the gene, (ii) the difference in resistance marker and (iii) the presence of a 6× histidine tag do not alter expression of the fluorescent proteins.

Finally, the system was challenged further by examining co-expression of the three different fluorescent proteins, each from a different vector. In total 39 (co-)expression experiments, including the controls, were performed where each pCoGW vector encoding an individual fluorescent protein was co-transformed with both combinations of the two pCo0GW vectors encoding the other fluorescent proteins (see Table 4 in Ref. [14]). Fig. 1C displays the images obtained for one combination and shows that the three