Available online at www.sciencedirect.com
ScienceDirect
Protein Expression and Purification 59 (2008) 232–241
Protein Expression & Purification
www.elsevier.com/locate/yprep

Development of a novel Gateway-based vector system for efficient, multiparallel protein expression in *Escherichia coli*

Felix Freuler *, Thomas Stettler, Marco Meyerhofer, Lukas Leder, Lorenz M. Mayr

Novartis Institutes for BioMedical Research (NIBR), Center for Proteomic Chemistry (CPC), Forum 1, Novartis Campus, CH-4056 Basel, Switzerland

Received 7 November 2007, and in revised form 4 February 2008
Available online 21 February 2008

Abstract

We describe a cloning and expression system which is based on the *Escherichia coli* T7 expression system and Gateway recombination technology. We have produced numerous destination vectors with selected fusion tags and an additional set of entry vectors containing the gene of interest and optional labeling tags. This powerful system enables us to transfer a cDNA to several expression vectors in parallel and combine them with various labeling tags. To remove the attached amino terminal tags along with the unwanted *att*B1 site, we inserted *PreScission* protease cleavage sites. In contrast to the commercially available destination vectors, our plasmids provide kanamycin resistance, which can be an advantage when expressing toxic proteins in *E. coli*. Some small-scale protein expression experiments are shown to demonstrate the usefulness of these novel Gateway vectors. In summary, this system has some benefits over the widely used and commercially available Gateway standard system, and it enables many different combinations for expression constructs from a single gene of interest.

Keywords: Fusion proteins; Protein expression; Protein labeling; Recombination cloning; Multiparallel expression

To date, many gene cloning and expression systems have been developed [1]. Usually the gene of interest (GOI)[1] is amplified and optionally modified by the polymerase chain reaction (PCR) and cloned into a plasmid, which in most cases is selected and replicated in *Escherichia coli*. The cloned plasmid can then either be used for protein expression in an appropriate host itself or as a shuttle vector (e.g. an entry vector for Invitrogen's Gateway system or a baculovirus transfer donor plasmid pFastBac1 for Invitrogen's Bac-to-Bac system).

While doing heterologous gene expression it has been found that not all hosts are suited for the production of a particular protein. This statement is especially true for *E. coli*, which is in many cases only useful for the expression of either distinct domains from larger proteins or small full-length proteins having a chain length of not more than a few hundred amino acids. In addition, the codon usage and potential regulatory sequences can negatively interfere with heterologous gene expression. For production of large eukaryotic multi-domain proteins or proteins which require proper post-translational modifications such as glycosylation, most scientists make use of more sophisticated expression systems such as yeast (e.g. *Pichia pastoris*), insect cells (e.g. *Spodoptera frugiperda* with or without baculovirus), and mammalian cells (e.g. HEK293, CHO) [2,3]. However, *E. coli* is still an attractive organism because it is easy to handle and generates only modest costs. One approach to overcome some expression issues with *E. coli* is the use of dedicated fusion proteins, which are usually attached to the amino terminus (N-terminus) of the target protein. Such a fusion protein partner (e.g. maltose binding protein of *E. coli*) can dramatically increase the expression level as well as the soluble fraction

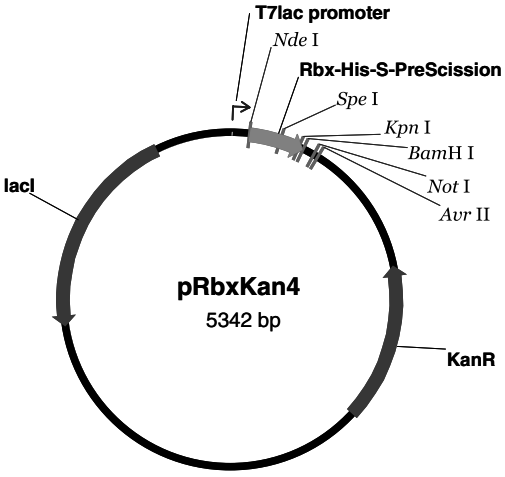
* Corresponding author. Address: Novartis Pharma AG, NIBR/CPC, WSJ-088.6.01A, CH-4002 Basel, Switzerland. Fax: +41 61 3245155.
E-mail address: felix.freuler@novartis.com (F. Freuler).

[1] *Abbreviations used:* GOI, gene of interest; PCR, polymerase chain reaction; Rbx, rubredoxin; GST, glutathione-*S*-transferase; Trx, thioredoxin; MBP, maltose binding protein; Ubi, ubiquitin.

1046-5928/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2008.02.003

of the heterologous gene product. Further, additional tags can be included into the construct for purification, detection and site specific modification of the target protein. For instance, we have successfully produced biotinylated proteins *in vitro* and *in vivo* by using the specific biotinylation tags BioEasy and AviTag [4].

Our expression vectors were originally based on pET vectors from Novagen (Merck Chemicals Ltd., UK). We modified the commercially available pET41a(+) by replacing the codons for the enterokinase cleavage site with those coding for the human rhinovirus 3C protease (PreScission from GE Healthcare Life Sciences). In addition, the thrombin protease recognition site was removed, and the amino terminal cassette containing GST-, His-, and S-tag was replaced by a synthetic DNA containing rubredoxin (Rbx) [5] and an adjacent His- and S-tag. The resulting plasmid, which was called pRbxKan4 (Fig. 1), was the base vector used to create a series of *E. coli* expression vectors carrying various amino terminal tags between the NdeI and the SpeI restriction sites. Additional vectors were made by substituting the rubredoxin tag with *Schistosoma japonicum* glutathione-*S*-transferase (GST), *E. coli* thioredoxin (Trx), nonsecreted maltose binding protein (MBP), human ubiquitin (Ubi), the IgG binding domain of *Streptococcus* sp. immunoglobulin G binding protein G (GB1), human SUMO1, and two IgG binding domains of *Staphylococcus aureus* immunoglobulin G binding protein A (ZZ). An overview of these tags is given in Table 1.




```
----T7 promoter---> ---------lac operator-------
        10         20         30         40         50         60         70
taatacgact cactataggg gaattgtgag cggataacaa ttcccctcta gaaataattt tgtttaactt
attatgctga gtgatatccc cttaacactc gcctattgtt aaggggagat ctttattaaa acaaattgaa

    -rbs--    NdeI ------------------------ rubredoxin ---------------------
        80       |  90         100        110        120        130        140
taagaaggag atatacatAT GAAAAAATAC CGTTGCAAAC TGTGCGGTTA CATCTACGAT CCGGAACAGG
attcttcctc tatatgtaTA CTTTTTTATG GCAACGTTTG ACACGCCAAT GTAGATGCTA GGCCTTGTCC
                   M   K  K  Y   R  C  K   L  C  G  Y   I  Y  D   P  E  Q

-------------------------------- rubredoxin --------------------------------
       150        160        170        180        190        200        210
GTGATCCGGA CTCCGGTATC GAACCGGGTA CTCCGTTCGA AGACCTGCCG GACGACTGGG TTTGCCCGCT
CACTAGGCCT GAGGCCATAG CTTGGCCCAT GAGGCAAGCT TCTGGACGGC CTGCTGACCC AAACGGGCGA
G  D  P  D   S  G  I   E  P  G   T  P  F  E   D  L  P   D  D  W   V  C  P  L

---------- rubredoxin ----------------- SpeI             ----- His-tag -----
       220        230        240       |250        260        270        280
GTGCGGTGCT TCCAAAGAAG ACTTCGAACC GGTTGAAACT AGTGGTTCCG GTCACCACCA CCACCACCAC
CACGCCACGA AGGTTTCTTC TGAAGCTTGG CCAACTTTGA TCACCAAGGC CAGTGGTGGT GGTGGTGGTG
  C  G  A   S  K  E   D  F  E  P   V  E  T   S  G  S   G  H  H  H   H  H  H

         ---------- S-tag ----------                    KpnI  -
       290        300        310        320        330        340     |  350
TCCGCTGGTA AAGAAACCGC TGCTGCTAAA TTCGAACGTC AGCACATGGA CTCTCCAGAC CTGGGTACCC
AGGCGACCAT TTCTTTGGCG ACGACGATTT AAGCTTGCAG TCGTGTACCT GAGAGGTCTG GACCCATGGG
 S  A  G   K  E  T  A   A  A  K   F  E  R   Q  H  M  D   S  P  D   L  G  T

------ PreScission ------BamHI ---- MCS --- NotI
       360        370     |                  |
TGGAAGTTCT GTTCCAGGGT CCAGGATCCN NNNNNNNNNN GCGGCCGC
ACCTTCAAGA CAAGGTCCCA GGTCCTAGGN NNNNNNNNNN CGCCGGCC
L  E  V  L   F  Q  G   P  G  S
```

Fig. 1. pRbxKan4. *E. coli* protein expression vector with N-terminal rubredoxin followed by a His-tag, an S-tag, and the *PreScission* cleavage site.