The pURI family of expression vectors: a versatile set of ligation independent cloning plasmids for producing recombinant His-fusion proteins

José Antonio Curiel [1], Blanca de las Rivas [1], José Miguel Mancheño [2], and Rosario Muñoz*[1]

Departamento de Microbiología, Instituto de Fermentaciones Industriales [1] and Grupo de Cristalografía y Biología Estructural, Instituto de Química-Física "Rocasolano" [2], CSIC, 28006-Madrid, Spain

* Corresponding author. Fax: +34-91-5644853.

E-mail address: rmunoz@ifi.csic.es (R. Muñoz)

ABSTRACT

A family of restriction enzyme–and ligation–independent cloning vectors has been developed for producing recombinant His–tagged fusion proteins in *Escherichia coli*. These are based on pURI2 and pURI3 expression vectors which have been previously used for the successful production of recombinant proteins at the milligram scale. The newly designed vectors combines two different promoters (lpp^{p}–5 and T7 RNA polymerase Ø10), two different endoprotease recognition sites for the His_6–tag removal (enterokinase and tobacco etch virus), different antibiotic selectable markers (ampicillin and erythromycin resistance), and different placements of the His_6–tag (N–and C–terminus). A single gene can be cloned and further expressed in the eight pURI vectors by using six nucleotide primers, avoiding the restriction enzyme and ligation steps. A unique *Not*I site was introduced to facilitate the selection of the recombinant plasmid. As a case study, the new vectors have been used to clone the gene coding for the phenolic acid decarboxylase from *Lactobacillus plantarum*. Interestingly, the obtained results revealed markedly different production levels of the target protein, emphasizing the relevance of the cloning strategy on soluble protein production yield. Efficient purification and tag removal steps showed that the affinity tag and the protease cleavage sites functioned properly. The novel family of pURI vectors designed for parallel cloning is a useful and versatile tool for the production and purification of a protein of interest.

Introduction

In many disciplines of protein research the use of recombinant proteins has increased greatly in recent years. Bacterial expression systems for heterologous protein production are attractive because of their ability to grow rapidly and at high density, low cost, and high productivity [1]. The efficient large–scale production of disparate pure and active recombinant proteins from different biological sources entails the optimization and standardization of several steps, namely, cloning and overexpression of the target gene, production of the soluble recombinant protein, purification, and (optionally) tag removal [2]. The importance of these aspects is revealed by the fact that there are numerous commercially and non–commercially available *E. coli* expression vectors that incorporate different backbones, promoters, fusion tags, and cloning procedures. Since sophisticated plasmid constructions are hardly engineered in a routine way with the exclusive use of specific endonuclease restriction sites, restriction enzyme–and ligation–independent cloning strategies of a target gene have become powerful tools [3, 4]. These strategies avoid the use of restriction endonuclease digestion and ligation of polymerase chain reaction products, allowing any gene to be cloned into an expression vector regardless of its sequence [5, 6].

Conversely, purification of recombinant proteins is accelerated by attachment of affinity tags to the N–or C–terminus of the expressed protein [7–10]. One of the simplest and most widely used tag is the His_6–tag which permits affinity purification by immobilized metal–affinity chromatography (IMAC). Despite its relative small size and charge of the His_6–tag, it is often useful to remove it for biological and functional studies since it can potentially interfere with the proper functioning of the protein [11]. Removal of the tag from a protein of interest can be accomplished with site–specific proteases, with the most

commonly used being enterokinase (EK), factor Xa, PreScission, tobacco etch virus (TEV), and thrombin. Nevertheless, even with the wide arsenal of expression vectors, tags and tag–removal strategies, structural genomics protein production facilities reveal success rates for soluble purified proteins less than 50% [12], which indicates that the choice of protein production strategy largely depends on the target protein [13].

In this scenario, we have previously described two expression vectors, pURI2 and pURI3, which enable parallel cloning of a given target gene for producing recombinant His–tagged proteins [5]. The vectors were constructed using the pT7–7 and pIN–III(lpp^{p}–5)A_3 plasmids as their template. At the N–terminus, the sequences of the His_6–tag and the enterokinase cleavage site were introduced in both vectors. They were designed to avoid the restriction enzyme and ligation steps during the cloning, so parallel cloning of the same polymerase chain reaction fragment can be carried out since both vectors shared the same leader sequence. As described elsewhere, we have successfully used these vectors to prepare protein samples that have been subjected to structural analyses by protein crystallography such as catabolic ornithine transcarbamylase from *Lactobacillus hilgardii* [5, 14], *p*–coumaric acid decarboxylase [15, 16] and a glycosidase [17] from *L. plantarum*. Moreover, several *L. plantarum* proteins overproduced by these vectors were biochemically characterized, a *p*–coumaric acid decarboxylase (PAD) [18], benzyl alcohol dehydrogenase [19], tannin acyl hydrolase [20], and nitroreductase [21], among others.

To increase the usefulness of pURI vectors and taking into account that the success in a protein production protocol depends on the specific protein, here we report the construction of new vectors from the family of pURI plasmids. These are also designed for an efficient purification of recombinant proteins. The pURI–TEV vectors incorporate a TEV cleavage site, the pURI–Ery vectors have erythromycin as selectable marker, and

finally, pURI–Cter vectors produce a His_6–tag in the C–terminus of the recombinant protein (Table 1).

Materials and methods

Strains

Escherichia coli **DH5αF´[F´/** *endA1 hsdR17*($r_k^-m_k^+$) *supE44 thi–1 recA1 gyrA*(Na1^r) **relA1 Δ(***lacIZYA–argF)U169 deoR (Φ80dlacΔ(lacZ)M15)*; Promega] was used for all DNA manipulations and for expression in pURI2–derived vectors. *E. coli* JM109 (DE3) [*endA1 recA1 gyrA96 hsdR17 supE44 relA1 thiΔ(lacpro)F´(traD36 proAB*$^+$ *lacI*^{q} *lacZΔM15)λc1857 ind1Sam/nin 5 lacUV5–T7* gene 1; Promega] was used for expression in pURI3–derived vectors. *Lactobacillus plantarum* CECT 748^T was purchased from the Spanish Type Culture Collection and *Lactobacillus brevis* RM 273 was isolated from wine [22]. Plasmids pIN–III(lpp^p–5)A3 [21, 22] and pT7–7 (USB) are expression vectors that allow the overexpression of the desired protein upon induction with isopropyl–β–D–thiogalactopyranoside (IPTG). The pUCE191 vector was constructed by cloning the fragment that contains the Ery^R gene from the pFB9 plasmid into the pUC19 vector [25]. Plasmid pFB9 is a plasmid originated from a Gram–positive host, *Staphylococcus aureus*, which was able to replicate and express Ery^R in a Gram–negative organism, *E. coli*. The plasmid–encoded Ery^R gene product N^6–dimethylates a specific adenine in 23S ribosomal RNA [26]. *E. coli* cells carrying pUCE191 could be grown either in erythromycin (200 μg/ml) or ampicillin (100 μg/ml). pUCE191 has been used for insertion–duplication

mutagenesis or illegitimate recombination between Gram–positive and Gram–negative species.

E. coli strains were cultured in Luria–Bertani (LB) medium [27] at 37 ºC and 200 rpm. When required, ampicilin or erythromycin was added to the medium at a concentration of 100 or 200 $\mu g \cdot mL^{-1}$, respectively. Chromosomal DNA, plasmid purification, and transformation of *E. coli* were carried out as described elsewhere [27].

Construction of pURI2–TEV and pURI3–TEV expression vectors

The construction of these new expression vectors was carried out by the same procedure used for the pURI2 and pURI3 construction [5] but replacing the sequence encoding the EK site by a TEV site in the forward oligonucleotide primer. The pURI2–TEV vector was constructed from plasmid pIN–III(lpp^{o}–5)A3 digested with *Xba*I and *Hin*dIII (Figure 1). The 230 bp insert was amplified by long primers 251 and 245 from *L. brevis* DNA, and digested with the same enzymes. The forward primer 251, from 5´to 3´direction, contained the following: 22 nucleotides that paired the pIN–III(lpp^{o}–5)A3 sequence from the unique *Xba*I recognition site, a His_6–tag, the TEV recognition site, and 35 nucleotides pairing the *L. brevis* sequence (251, 5´–
ACTCTAGAGGGTATTAATAATGGGGGGTTCT*CATCATCATCATCATCATCAT*GGTGAA
AACCTGTATTTCCAGGGCATGCCTGCTACTGCTAATCGCTATCATTTTGGCGG)
(the *Xba*I restriction site is underlined, the sequence coding the TEV recognition site is written in bold, and that coding for the poly–His tag is indicated in italics). The reverse primer, primer 245, is the one used previously for the construction of pURI2 vector [5]. The

amplified fragment was digested with the *Xba*I and *Hin*dIII enzymes, and subsequently ligated to the pIN–III(lpp^{p}–5)A3 digested with the same enzymes.

Expression vector pURI3–TEV was constructed by a similar protocol (Figure 2). Expression vector pT7–7 was digested with *Nde*I and *Hin*dIII restriction enzymes. The *L. brevis* internal fragment was amplified by long oligonucleotides 252 and 245. Oligonucleotide 252 was identical to primer 251, except that the nucleotides based on the pIN–III(lpp^{p}–5)A3 sequence were substituted by nucleotides based on the pT7–7 sequence containing a *Nde*I restriction site (252, 5´–AGATATACATATGGGGGGTTCT*CATCATCATCATCATCAT*GGTGAAAACCTGTATTTCCAGGGCATGCCTGCTACTGCTAATCGCTATCATTTTGGCGG) (the *Nde*I restriction site is underlined, the TEV recognition site is written in bold, and the His_6–tag is indicated in italics). The amplified fragment was digested with *Nde*I and *Hin*dIII enzymes and ligated to the vector pT7–7 digested with the same enzymes.

Sequences of the pURI2–TEV and pURI3–TEV vectors were verified by restriction analysis and DNA sequencing.

Construction of pURI2–Ery and pURI3–Ery expression vectors

The construction of these new expression vectors was carried out by the restriction enzyme– and ligation–independent cloning strategy used to clone any gene into pURI vectors [5]. The erythromycin resistance gene, 732 bp, was amplified from the pUCE191 vector [25] by *Pfu* DNA polymerase using the following oligonucleotides, forward 697 (5´–*CAATAATATTGAAAAAGGAAGAGTatg*aacgagaaaaatataaaacaca) and reverse 698 (5´–

*ATGAGTAAACTTGGTCTGACAGtta*cttattaaataatttatagctatt) (the nucleotides pairing the expression pURI vector sequence are indicated in italics, and the nucleotides pairing the erythromycin gene sequence are written in lowercase letters; the start and stop codons are indicated in bold). After the amplification, the 732 bp fragment was used to substitute the ampicillin resistance gene present in pURI vectors, by using this fragment as homologous primer pair in a PCR reaction using pURI2 or pURI3 vectors as template. The product of this PCR was digested with *Dpn*I that exclusively restricts methylated DNA. *E. coli* cells were transformed directly with the digestion product and plated on LB plates containing erythromycin at 200 μg/ml. The construction of the pURI2–Ery and pURI3–Ery vectors was verified by DNA sequencing (Figure 3).

Construction of pURI2–Cter and pURI3–Cter expression vectors

The construction of these pURI vectors containing the His_6–tag placed on the C–terminus of the recombinant protein was carried out by the strategy used for the construction of pURI2 and pURI3 vectors [5]. The pURI2–Cter vector was constructed from plasmid pIN–III(lpp^p–5)A3 digested with *Xba*I and *Hin*dIII (Figure 4). The 230 bp insert was amplified from *L. brevis* DNA by the primers 412 and the long primer 411, and digested with the same enzymes. The forward primer 412, from 5´to 3´direction, contained 22 nucleotides that paired the pIN–III(lpp^p–5)A3 sequence from the unique *Xba*I recognition site and 22 nucleotides pairing the *L. brevis* sequence (412, 5´–ACTCTAGAGGGTATTAATAATGcctgctactgctaatcgctatc) (the *Xba*I restriction site is underlined, and the nucleotides pairing the *L. brevis* sequence are written in lowercase

letters). The reverse long primer, primer 411, contained from 5´to 3´direction, a *Hin*dIII restriction site, several stop codons arranged in tandem, a His_6–tag, a *Not*I restriction site, and 27 nucleotides pairing the *L. brevis* sequence (411 , 5´–CCAAGCTTAGTTAGCTATTA*ATGATGATGATGATGATG*TGCGTAGCGGCCGCAGATTCAGTAAAGCCTCGTGTCGCTCG) (the *Hin*dIII and *Not*I restriction sites are underlined, and the poly–His tag is written in italics). The amplified fragment was digested with *Xba*I and *Hin*dIII and ligated to the pIN–III(lpp^o–5)A3 cut with the same enzymes.

Vector pURI3–Cter was constructed by the same strategy (Figure 5). The *L. brevis* DNA fragment was amplified by the primer 410 and the long primer 411, which had been explained above. The forward primer 410, from 5´to 3´direction contained an *Nde*I restriction site followed by 22 nucleotides pairing the *L. brevis* sequence (410, 5´–AGATATACATATGcctgctactgctaatcgctatc) (the *Nde*I restriction site is underlined, and the nucleotides pairing the *L. brevis* sequence are written in lowercase letters). The amplified DNA fragment was digested with *Nde*I and *Hin*dIII restriction enzymes, and ligated to the pT7–7 vector digested with the same enzymes.

Restriction analysis and DNA sequencing was used to verify the pURI2–Cter and pURI3–Cter sequences.

Cloning and expression of the target gene and purification of the encoded protein

The gene encoding the phenolic acid decarboxylase (PAD) from *L. plantarum* was cloned in all the pURI vectors to check the efficacy of the new vectors. The aim of the vector family is to allow the use of only few (six) oligonucleotides for the cloning of a

specific gene in the eight different pURI vectors (Table 2). The cloning of the PAD gene in pURI3 vector has been previously described [15, 18]. The gene coding for the *L. plantarum* PAD was cloned into pURI3 vector by using 274 and 275 oligonucleotides [15, 18], being primer 274 (5´–CATCATGGTGACGATGACGATAAGatgacaaaaactttaaaacacttg) and 275 (5´–AAGCTTAGTTAGCTATTATGCGTAttacttatttaaacgatggtagttt). The same oligonucleotide pair is used to clone into pURI2, pURI2–Ery and pURI3–Ery vectors. To clone into pURI2–TEV and pURI3–TEV vectors, oligonucleotides 413 and 275 were used (primer 413, 5´–GGTGAAAACCTGTATTTCCAGGGCatgacaaaaactttaaaacacttg). Finally, to clone into pURI2–Cter, the *pad* gene was amplified by 419 and 420 primer set, and oligonucleotides 421 and 420 were used to clone into pURI3–Cter vector (primer 419, 5´–CTGGAACTCTAGAGGGTATTAATAatgacaaaaactttaaaacacttg; 420, 5´–GCTATTA*ATGATGATGATGATGATG*cttatttaaacgatggtagttttg; 421, 5´–TAACTTTAAGAAGGAGATATACATatgacaaaaactttaaaacacttg) (the nucleotides pairing the expression vector sequence are indicated in capital letters and the nucleotides pairing the *pad* gene sequence are written in lower case; the sequence coding the His_6–tag is indicated in italics, and the start and stop codons of the recombinant protein are written in bold). All the PCR amplifications were for 30 cycles with the following conditions: 95 ºC, 30s; 55 ºC, 1 min; 72 ºC, 1 min. After amplification, the 0.5 kb PCR products were gel–purified and inserted into pURI vectors by using the enzyme restriction– and ligation–free cloning strategy described previously for pURI2 and pURI3 vectors [5]. Briefly, the purified PCR products were used as homologous primer pair in a PCR reaction using the corresponding expression vectors as template. The product of this PCR was digested with *Dpn*I, that exclusively restrict methylated DNA and, later, with *Not*I, an enzyme that only

cuts the original copies of pURI vectors. *E. coli* **DH5α cells were t**ransformed directly with the digestion product [5].

pURI2–derivatives were amplified in the *E. coli* **DH5α** or JM109 (DE3) cells, and pURI3–derivatives plasmids in the host JM109 (DE3) *E. coli* strain. Ten different colonies were selected from *E. coli* JM109 (DE3) (pURI2—TEV–PAD) to observe cell–to–cell variation on protein production. Cells carrying a recombinant plasmid were grown at 37 ºC in Luria–**Bertani media containing ampicillin (100 μg/mL final concentration) or erythromycin (200 μg/mL final concentratio**n), when required, and induced by adding IPTG (0.4 mM final concentration) when culture OD_{600nm} was 0.5. After induction, the cells were grown at 22 ºC during 20 h and collected by centrifugation. Crude extracts were prepared by French Press lysis (three cycles at 1100 psi) of cell suspensions obtained by suspending the frozen cell paste with 20 mM Tris–HCl, pH 8.0, 100mM NaCl. The insoluble fraction of the lysate was removed by centrifugation at 47000*g* for 30 min at 4 ºC. The relative proportions of soluble and insoluble material that were expressed for each of the constructs were determined. Samples of both the soluble and insoluble fractions were analyzed by SDS–PAGE. The presence of PAD protein on the gels was quantified by using the ChemiDoc XRS+ Imagin System (Bio–Rad) using pURI2/pURI3 as controls.

The soluble cell homogenate, obtained by French Press lysis of the *E. coli* cells containing the *pad* gene cloned on each one of the pURI vectors, was then applied to a His–Trap–FF crude chelating affinity column (Amersham Biosciences, Uppsala, Sweden) using an ÄKTA–Prime system (Amersham Pharmacia Biotech). The column was equilibrated with 20 mM Tris–HCl, pH 8.0, 100 mM NaCl containing 10 mM imidazole, to improve the interaction specificity in the affinity chromatography step. The bound enzyme was eluted by applying a continuous gradient of imidazole concentration, from 20 mM Tris–HCl, pH

8.0, 100 mM NaCl containing 10 mM imidazole to 20 mM Tris–HCl, pH 8.0, 100 mM NaCl containing 500 mM imidazole. Fractions containing the eluted PAD were pooled and the protein was then dialyzed against 20 mM Tris–HCl, pH 8.0, 100 mM NaCl. The concentration of PAD was estimated using the Bio–Rad protein assay. In the fused TEV–PAD protein produced by the pURI3–TEV vector, the PAD protein (5 mg/mL) was incubated with one unit of recombinant TEV at 4 ºC for 20 h. After digestion, the sample was run on a sodium dodecyl sulfate (SDS) gel to check the cleavage.

PAD activity was determined by following the decarboxylation of *p*–coumaric acid into vinyl phenol according to the method previously described [18]. Briefly, the standard reaction was performed by adding 200 µg/mL of protein from the cell–free extract or adding 3 µg/mL of purified PAD enzyme into a solution containing 1mM *p*–coumaric acid in 25 mM phosphate buffer (pH 6.5). The reaction was incubated at 37 ºC during 1h, for cell extracts, or 20 min for pure PAD protein. The reaction products were extracted twice with ethyl acetate and analyzed by HPLC.

Results and Discussion

Characterization of the pURI Family Expression Vectors

One of the main conclusions derived from structural genomics protein production facilities is that the success rates for soluble purified proteins largely depends on the target protein itself [13] and therefore the choice of the adequate system for production of a

specific protein is mainly based on a trial–and–error approach. With the aim of providing a new tool for the production of recombinant proteins, we previously created two expression vectors pUIR2 and pURI3 based on the pIN–III(lpp^p–5)A3 and pT7–7 plasmids, respectively, which share the same cloning procedure and protein purification setup [5]. These vectors were specifically designed to avoid the restriction enzyme and ligation steps during the cloning, and included an N–terminal His_6–tag which allows convenient affinity purification of the target protein from crude cell extracts and also an enterokinase cleavage site that leaves no extra residues in the protein upon tag removal. In addition, a unique *Not*I site was introduced to facilitate the selection of the adequate recombinant plasmid. Moreover, pURI2 and pUIR3 vectors included a 230 bp *L. brevis* noncoding intergenic region to facilitate the integration of long DNA fragments. Parallel cloning of the same PCR fragment can be carried out since both vectors shared the same leader sequence. These vectors were successfully used to hyperproduce several proteins from lactic acid bacteria [5, 14–21].

To increase the usefulness of the pURI vectors we decided to expand its versatility by creating a vector family, which combines different protease recognition sites, different antibiotic selectable markers, and also different placement of the His_6–tag (Table 1). Two endoproteases were considered for tag removal namely, enterokinase and TEV. Enterokinase specifically recognizes a five–amino acid polypeptide (DDDKX) and cleaves at the carboxyl site of the lysine side chain. However, sporadic cleavage at other residues was observed to occur at low levels, depending on the conformation of the protein [28]. Conversely, TEV protease is also an ideal tool for removing tags from fusion proteins due to high specificity; the optimal recognition sequence is ENLYFQG, with the cleavage occurring between the conserved glutamine and glycine residues. TEV protease is attractive

because it can be expressed and purified in the lab (making it cost–effective), and leaves only one glycine residue at the N–terminus of the protein upon cleavage [30]. The optimal temperature for cleavage is 34 ºC; however, the enzyme can also be used at temperatures as low as 4 ºC and is active under a wide range of conditions and in the presence of various protease inhibitors [31–33]. To construct pURI2–TEV and pURI3–TEV vectors, long oligonucleotides pairing the *L. brevis* sequence and containing all the required elements (His_6–tag, TEV recognition site) were used as described in Materials and Methods. Expression vector pURI2–TEV was constructed by using 251 and 245 primers, and vector pURI3–TEV by using 252 and 245 primers. Both expression vectors shared a common 316 bp fragment from the ATG start codon upstream from the His_6–tag coding sequence to the *Hin*dIII restriction site (Figure 1 and 2). The common DNA fragment is identical to the 310 pb fragment included in the pURI2/pURI3 vectors but replacing the enterokinase recognition site by the TEV protease site (Table 1).

The selective markers "amp" (ampicillin resistance, also abbreviated Ap or *bla* for β–lactamase) and "Ery" (erythromycin resistance, also abbreviated em) have been widely used. While ampicillin resistance is commonly used for selection in a variety of cloning vectors, erythromycin resistance may be preferable under certain specific conditions, and when subcloning target genes from other ampicillin–resistant vectors. The erythromycin resistance gene from pUCE191 was amplified with oligonucleotides 697 and 698. The 732 bp amplified fragment was used to replace the ampicillin resistance gene, 858 bp, found in the pURI2 and pURI3 vectors. The 732 erythomycin fragment was introduced into pURI2 and pURI3 vectors by the restriction enzyme– and ligation–independent cloning strategy used to clone any gene into the pURI vector family. The new vectors, pURI2–Ery and pURI3–Ery will permit the production of recombinant proteins with the same characteristics

as those produced by the formers pURI2 and pURI3 vectors (e.g., initial methionine followed by a three amino acid spacer, a six histidine affinity tag, a spacer glycine residue, and the five amino acid enterokinase recognition site) (Figure 3, Table 1). The direction of the transcription of the drug resistance gene also needs to be considered. In all pURI vectors, the resistance genes Amp and Ery are in opposite orientation from the lpp^{p}–5 (pURI2) or Ø10 (pURI3) promoters, so induction of these promoters should not result in an increase in antibiotic resistance gene product. In order to increase the versatility of the pURI vectors, in addition to ampicillin and erythromycin resistance, kanamycin resistance pURI–derivatives have been also constructed by using the same strategy (data not shown)

Poly–histidine affinity tags are placed on either the N–or the C–terminus of recombinant proteins, although the optimal location does vary depending on the folding and biochemical characteristics of the target protein. In principle, it cannot be excluded that the affinity tag may interfere with protein activity, although the relative small size and charge of the His_6–tag ensure that protein activity is rarely affected. Halliwell et al. (2001) [34] fused a poly–His tag to either the N–or the C–terminus of L–lactate dehydrogenase of *Bacillus stearothermophilus*, and found that the C–terminally tagged enzyme displayed lower activity compared both to the wild–type and to the N–terminally tagged variant, probably due to a misfolding of the enzyme. In enzymes possessing a three–dimensional structure available, it is necessary to check whether the N–terminus or the C–terminus are noninteracting with the substrate and cofactor–binding sites. In order to offer production of recombinant C–terminally tagged proteins, pURI2–Cter and pURI3–Cter expression vectors were created (Figure 4 and 5, Table 1). To construct pURI2–Cter vector oligonucleotides 412 and 411 were used; similarly, pURI3–Cter was constructed by using 410 and 411 oligonucleotides. Both expression vectors shared a common 280 bp fragment from the

ATG start codon to the *Hin*dIII restriction site downstream from the His_6–tag coding sequence. The common DNA fragment encodes for the following: a N–terminal methionine followed by the 230 bp noncoding intergenic *L. brevis* sequence, a rare–cut *Not*I restriction site, a six histidine affinity tag, four stop codons arranged in tandem, and finally, a *Hin*dIII restriction site (Figure 4 and 5).

As indicated in Table 2, by the use of only six nucleotide primers, a gene can be expressed in eight different expression vectors by using a similar restriction enzyme– and ligation–independent cloning strategy for producing a recombinant His–fusion protein. The designed pURI vectors allows to use different promoters (*lpp* and T7 RNA polymerase promoters), different protease cleavage sites (enterokinase and TEV protease), different selective markers (ampicillin and erythromycin) and different poly–His tag placements (N– or C–terminus) for the production of a protein of interest (Table 1).

Cloning and expression of the gene encoding the L. plantarum PAD protein and purification of the recombinant protein

PAD enzymes are biotechnologically relevant enzymes that catalyze the conversion of ferulic or *p*–coumaric acids into the corresponding volatile derivatives 4–vinyl guaiacol (3–methoxy–4–hydroxystyrene) or 4–vinyl phenol (4–hydroxystyrene), considered as precursors of vanillin (4–hydroxy–3–methoxybenzaldehyde) production. Vanillin is the most commonly used flavour in foods, beverages, perfumes or pharmaceuticals. Vinyl guaiacol is priced around 40 times more than ferulic acid, and it can be biotransformed further to acetovanillone, ethylguaiacol, and vanillin [35]. As a styrene–type molecule, vinyl guaiacol

can be polymerized; the resultant oligomer [poly(3–methoxy–4–hydroxystyrene)] was found to be easily biodegradable [36]. In addition, the 4–vinyl guaiacol or 4–vinyl phenol, produced by the activity of PAD enzymes on hidroxycinnamic acids, are considered to be food additives and are approved as flavouring agents by regulatory agencies [37].

By using the pURI3 plasmid, we have overproduced, crystallized and solved the three dimensional structure of PAD from *L. plantarum* [15, 16, 18]. The procedure employed for cloning the *pad* gene into pURI3 vector can also be used to clone into each one of the vectors from the pURI family. Briefly, PCR products of the gene of interest need to be generated with specific overhangs that are complementary to the integration site sequence of the vectors. The nucleotide sequence of the gene encoding the target protein is amplified by PCR using specific oligonucleotides (Table 2) and *Pfu* DNA polymerase. Generally, the sequence of forward primers include the sequence showed in Table 2, where the last three nucleotides encode the first methionine residue of the target protein, followed by nucleotides pairing the sequence from the second amino acid residue of the target protein. Similarly, reverse primers include the sequence showed in Table 2, followed by nucleotides pairing the sequence up to the stop codon of the target protein; however, to clone in the pURI–Cter vectors, the protein stop codon is not included into the primer sequence, as it is included after the His_6–tag (Table 2). The PCR fragment is added as megaprimer to the methylated recipient template plasmid. After the denaturation step, the homologous parts of the PCR fragment hybridize to the defined integration site and these fragments are elongated in vitro by the *Pfu* DNA polymerase. The elongated and modified strands are not methylated. When PCR is finished, they are treated with the *Dpn*I enzyme, which exclusively restricts methylated DNA, removing the methylated templates. An

additional screening tool herein used is a *Not*I digestion of the template plasmid, which can only be used if the target gene sequence lacks such restriction site.

The pURI2–derived expression vectors could be used in a single strain of *E. coli* in all steps from plasmid construction to the expression of the target gene. However, in pURI3–derived vectors, the plasmid needs to be cloned into any cloning *E. coli* strain, and for expression, into *E. coli* DE3 cells. Therefore, *E. coli* DE3 cells could be used for gene expression in all the designed pURI vectors. The gene coding for the PAD protein was expressed into all the pURI family vectors by IPTG induction (0.4 mM final concentration). Following expression, the cell densities of the cultures were normalized to an OD_{600nm} of 5.0 prior to lysis, which allowed for a uniform number of cells for all the samples. The cultures were then lysed and centrifuged to sediment insoluble cell debris. The six new pURI vectors were checked for PAD production, and the relative proportions of soluble and insoluble expression of PAD was assessed. Two important aspects can be deduced from Figure 6A: firstly, soluble PAD is detected in all the cultures and secondly, the protein production levels were dependent on the pURI vector used. Taking into account the size of the His_6–tag, the expected size of the hyperproduced protein is 21.8 kDa. An obvious soluble protein production from crude extracts was obtained by pURI–Cter vectors, especially by pURI3–Cter. Even though no apparent PAD hyperproduction was achieved by using other pURI vectors (e.g., pURI2–TEV) (Figure 6 and 7), PAD proteins were purified from all the extracts by IMAC (data not shown). After purification, the final protein production yield varied from 1.29 mg/l (pURI2–TEV) to 13.84 mg/l (pURI3–Cter). As shown in Figure 6B, PAD was largely expressed as insoluble material with in Cter derived vectors (pURI2–Cter and pURI3–Cter), although soluble protein was largely obtained in both vectors. In contrast, pURI2 and pURI3 vectors rendered the highest proportion of

soluble PAD protein, which produced PAD in the proportions of 4:1 and 6:1 (soluble:insoluble), respectively (Figures 6 and 7). Additionally, Figure 7 indicates that both pURI–TEV vectors produced the highest yield of the insoluble form of the protein. In order to check cell–to–cell variation on protein production, ten different *E. coli* JM109 (DE3) cells transformed with pURI2–TEV–PAD vector were selected. Soluble cell–extracts from cultures of these colonies revealed that PAD production was similar in all of them, sharing a common expression profile (Figure 9). This result indicated that the different protein yield observed among the pURI vectors are due to vector borne enhancement and not to clonal variances.

As *E. coli* DE3 cells could be used for gene expression in all the pURI vectors, these cells were used to produce PAD by using each one of the eight designed vectors (Figure 10). The protein production profile showed by the different pURI2 vectors on *E. coli* JM109 (DE3) cells was similar to that showed on *E. coli* **DH5α. Vector pURI2**–TEV–PAD yielded the lowest soluble protein production, whereas pURI2–Cter–PAD produced the highest PAD yield (48.45 mg/l), more than three times **higher than on DH5α cells** (Figure 6). Therefore, the choice of the *E .coli* cells used as a host had a relevant role in protein production [1].

Therefore, this study illustrates the usefulness of the new family of pURI expression vectors when considered as tool for the parallel cloning and expression of a recombinant protein whose production yield may vary significantly on the vector used.

The activity of the purified His_6–tagged PAD proteins was determined by measuring the decarboxylation of *p*–coumaric acid into vinyl phenol. When cell extracts were assayed the reaction was fully decarboxylated to vinyl phenol, except by pURI2–Ery and pURI2–TEV extracts which transformed 90% or only a 20% of the substrate. Both cell extracts

contained the lowest PAD protein in the soluble form (Figure 7A). However, when the different purified PAD proteins were compared, all showed the same specific activity (data not shown), indicating that the location of the His_6–tag as well as the protease cleavage site did not affect catalytic activity.

The proper function of the TEV cleavage site was demonstrated on the recombinant PAD protein produced in the pURI3–TEV vector. The fused PAD protein was purified, dialyzed, and digested with TEV protease at 4 ºC. Figure 8 showed the digested PAD sample, it could be observed that the suppression of the leader sequence containing the affinity tag (17 amino acid residues) produces an expected reduction in the protein size, from a predicted 23 kDa (196 amino acid fused protein) to a 21 kDa cleaved protein (179 amino acids). The functionality of the EK cleavage site on pURI vectors was previously demonstrated [5].

In summary, since the production yield of soluble recombinant protein largely depends on the target protein itself, and considering that one of the major bottlenecks in the structural studies is the large–scale production of pure and active proteins, we herein provide a new family of expression vectors that may allow the overproduction and further purification of soluble recombinant proteins. These vectors are based on the previously described and successfully used pURI2 and pURI3 plasmids. The usefulness of the pURI vectors was increased by the presence of different protease recognition sites to remove the His_6–tag, different antibiotic selectable marker, and different placement of the affinity tag. By the use of only six nucleotide primers a gene can be expressed in eight different expression vectors with a similar restriction enzyme–and ligation–independent cloning strategy. The target protein chosen for the validation of the pURI vector family has evidenced the different protein production obtained by each one of the pURI vectors, and

also pointed out the difficulty to decide empirically which system is the best for a heterologous protein production as it depends often on the target protein itself. Therefore the availability of a set of different, but related, expression vectors for parallel cloning confers great advantages for the production of a protein of interest.

Acknowledgments

This work was supported by grants RM2008-00002 (Instituto Nacional de Investigación Agraria y Alimentaría), AGL2008-01052, Consolider INGENIO 2010 CSD2007-00063 FUN-C-FOOD (Comisión Interministerial de Ciencia y Tecnología), and S-0505/AGR/000153 and S2009/AGR-1469 (ALIBIRD) (Comunidad de Madrid). JMM thanks the *Ministerio de Ciencia e Innovación* for a research grant (BFU2007-67404/BMC) and "Factoría de Cristalización" Consolider-Ingenio 2010 in support of his research. The technical assistance of M.V. Santamaría is greatly appreciated. J.A Curiel is a recipient of a predoctoral fellowship from the MEC. The pURI family vectors are available from the authors upon request.

References

[1] K. Terpe, Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundaments to commercial systems, Appl. Microbiol. Biotechnol. 72 (2006) 211-222.

[2] I. Hunt, From gene to protein: a review of new and enabling technologies for multi–parallel protein expression, Prot. Exp. Purif. 40 (2005) 1–22.

[3] D. Tillet, B.A. Neilan, Enzyme free cloning: a rapid method to clone PCR products independent of vectors restriction enzyme sites, Nucleic Acids Res. 27 (1999) e29.

[4] H. Tseng, DNA cloning without restriction enzyme and ligase, BioTechniques 27 (1999) 1240–1244.

[5] B. De las Rivas, J.A. Curiel, J.M. Mancheño, R. Muñoz, Expression vectors for enzyme restriction–and ligation–independent cloning for producing recombinant His–fusion proteins, Biotechnol. Prog. 23 (2007) 680–686.

[6] L. Stols, M. Gu, L. Dieckman, R. Raffen, F.R. Collart, M.I. Donnelly, A new vector for high–throughput, ligation–independent cloning encoding a tobacco etch virus protease cleavage site, Prot. Expr. Purif. 25 (2002) 8–15.

[7] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, Appl. Microbiol. Biotechnol. 60 (2003) 523–533.

[8] J.J. Lichty, J.L. Malecki, H. D. Agnew, D.J. Michelson–Horowitz, S. Tan, Comparison of affinity tags for protein purification, Prot. Expr. Purif. 41 (2005) 98–105.

[9] D.S. Waugh, Making the most of affinity tags, Trends Biotech. 23 (2005) 316–320.

[10] D. Esposito, D. K. Chatterjee, Enhancement of soluble protein expression through the use of fusion tags, Curr. Opin. Biotech. 17 (2006) 353–358

[11] J. Arnau, C. Lauritzen, G.E. Petersen, J. Pedersen, Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins, Prot. Expr. Purif. 48 (2006) 1–13

[12] Structural Genomics Consortium, China Structural Genomics Consortium, Northeast Structural Genomics Consortium, Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., et al., Protein production and purification, Nature Methods 5 (2008) 135–146

[13] A. Malhotra, Tagging for protein expression, Meth. Enzymol. 463 (2009) 239–258

[14] B. de la Rivas, G.C. Fox, I. Angulo, M.M. Ripoll, H. Rodríguez, R. Muñoz, J.M. Mancheño, Crystal structure of the hexameric catabolic ornithine transcarbamylase from *Lactobacillus hilgardii*: Structural insights into the oligomeric assembly and metal binding, J. Mol. Biol. 393 (2009) 425–434.

[15] H. Rodríguez, B. de las Rivas, R. Muñoz, J.M. Mancheño, Overexpression, purification, crystallization and preliminary structural studies of *p*-coumaric acid decarboxylase from *Lactobacillus plantarum*, Acta Cryst. F63 (2007) 300–303.

[16] H. Rodriguez, I. Angulo, B. de las Rivas, N. Campillo, J.A. Páez, R. Muñoz, J.M. Mancheño, *p*-Coumaric acid decarboxylase from *Lactobacillus plantarum*: structural insights into the active site and decarboxylation catalytic mechanism, Proteins 78 (2010) 1662–1676.

[17] I. Acebrón, J.A. Curiel, B. de las Rivas, R. Muñoz, J.M. Mancheño, Cloning, production, purification and preliminary crystallographic analysis of a glycosidase from the food lactic acid bacterium *Lactobacillus plantarum* CECT 748^{T}, Prot. Expr. Purif. 68 (2009) 177–182.

[18] H. Rodríguez, J.M. Landete, J.A. Curiel, B. de las Rivas, J.M. Mancheño, R. Muñoz, Characterization of the *p*-coumaric acid decarboxylase from *Lactobacillus plantarum* CECT 748^{T}. J. Agric. Food Chem. 56 (2008) 3068–3072.

[19] J. M. Landete, H. Rodríguez, B. de las Rivas, R. Muñoz R, Characterization of a benzyl alcohol dehydrogenase from *Lactobacillus plantarum* WCFS1, J. Agric. Food Chem. 56 (2008) 4497–4503.

[20] J.A. Curiel, H. Rodríguez, I. Acebrón, J. M. Mancheño, B. de las Rivas, R. Muñoz, Production and physicochemical properties of recombinant *Lactobacillus plantarum* tannase, J. Agric. Food Chem. 57 (2009) 6224–6230.

[21] H. Guillén, J.A. Curiel, J.M. Landete, R. Muñoz, T. Herraiz, Characterization of a nitroreductase with selective nitroreduction properties in the food and intestinal lactic acid bacterium *Lactobacillus plantarum* WCFS1, J. Agric. Food Chem. 57 (2009) 10457–10465.

[22] M.V. Moreno–Arribas, M.C. Polo, F. Jorganes, R. Muñoz, Screening of biogenic amine production by lactic acid bacteria isolated from grape must and wine, Int. J. Food Microbiol. 84 (2003) 117–123.

[23] Y. Masui, T. Mizuno, M. Inouye, Novel high level expression cloning vehicles: 10^4–fold amplification of *Escherichia coli* minor protein, Bio/Technology 2 (1984) 81–85.

[24] S. Inouye, M. Inouye, Up–promoter mutations in the *lpp* gene of *Escherichia coli*, Nucleic Acids Res. 13 (1985) 3101–3109.

[25] C. Arrecubieta, E. García, R. López, Sequence and transcriptional analysis of a DNA region involved in the production of capsular polysaccharide in *Streptococcus pneumoniae* type 3, Gene 167 (1995) 1–7.

[26] F. Barany, J.D. Boeke, A. Tomasz, Staphylococcal plasmids that replicate and express erythromycin resistance in both *Streptococcus pneumoniae* and *Escherichia coli*, Proc. Natl. Acad. Sci. USA. 79 (1982) 2991–2995.

[27] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular cloning: A Laboratory Manual, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.

[28] S.I. Choi, H.W. Song, J.W. Moon, B.L. Seong, Recombinant enterokinase light chain with affinity tag: expression from *Saccharomyces cerevisiae* and its utilities in fussion protein technology, Biotechnol. Bioeng. 75 (2001) 718–724.

[29] T.D. Parks, K.K. Leuther, E.D. Howard, S.A. Johnston, W.C. Dougherty, Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase, Anal. Biochem. 216 (1994) 413–417.

[30] B.B. Suh–Lailam, J.M. Hevel, Efficient cleavage of problematic tobacco etch virus (TEV)–protein arginine methyltransferase constructs, Anal. Biochem. 387 (2009) 130–132.

[31] J.A. Daros, M.C. Schaad, J.C. Carrington, Functional analysis of the interaction between VPg–proteinase (NIa) and RNA polymerase (NIb) of tobacco etch virus using conditional and suppressor mutants, J. Virol. 73 (1999) 8732–8740.

[32] J. Phan, A. Zdanov, A.G. Evdokimov, J.E. Tropea, H.K. Peters, K.B. Kapust, M. Li, A. Wlodawer, D.S. Waugh, Structural basis for the substrate specificity of tobacco etch virus protease, J. Biol. Chem. 277 (2002) 50564–50572.

[33] A.K. Mohanty, C.R. Simmons, M.C. Wiener, Inhibition of tobacco etch virus protease activity by detergents, Protein Expr. Purif. 27 (2003) 109–111.

[34] C.M. Halliwell, G. Morgan, C.P. Ou, A.E. Cass, Introduction of a (poly)histidine tag in L–lactate dehydrogenase produces a mixture of active and inactive molecules, Anal. Biochem. 295 (2001) 257–261.

[35] J.P.N. Rosazza, Z. Huang, L. Dostal, T. Volm, B. Rousseau, Review: Biocatalytic transformation of ferulic acid: An abundant aromatic natural product, J. Ind. Microbiol. 15 (1995) 457–471.

[36] H. Hatakeyama, E. Hayashi, T. Haraguchi, Biodegradation of poly (3–methoxy–4–hydroxystyrene), Polym. 18 (1977) 759–763.

[37] Joint Expert Committee on Food Additives. Evaluation of certain food additives and contaminants. Fifty–fith report of the Joint WHO/FAO Expert Committee on Food Additives. WHO Technical report series no. 901, World Health Organization, Geneva, 2001.

Figure captions

Fig. 1. (A) Schematic representation of the expression vector pURI2–TEV (7791 bp). (B) Sequence of pURI2–TEV containing the leader sequence encoding region. The expression leader sequence consists of a N–terminal methionine residue followed by a three amino acid spacer, a six histidine affinity tag, a spacer glycine residue, and the seven amino acid TEV recognition site, ENLYFQG, followed by an initial protein methionine. The expression leader sequences are under the control of the lpp^{p}–5 and lac^{PO} promoters. The TG motif typical of extended –10 boxes is also indicated. Cleavage of expressed protein by TEV protease occurs between the glutamine and glycine residues (QG). Restriction sites are indicated and underlined in the nucleotide sequence: X, *Xba*I, N, *Not*I, H, *Hin*dIII, and B, *Bam*HI. Tandem stop codons are indicated by asterisks.

Fig. 2. (A) Schematic representation of the expression vector pURI3–TEV (2696 bp). (B) Sequence of pURI3–TEV containing the leader sequence encoding region. The expression leader sequence consists of a N–terminal methionine residue followed by a three amino acid spacer, a six histidine affinity tag, a spacer glycine residue, and the seven amino acid TEV recognition site, ENLYFQG, followed by an initial protein methionine. The expression leader sequences are under the control of the T7 polymerase Ø promoter. Cleavage of expressed protein by TEV protease occurs between the glutamine and glycine residues (QG). Restriction sites are indicated and underlined in the nucleotide sequence: Nd, *Nde*I, N, *Not*I, H, *Hin*dIII, and C, *Cla*I. Tandem stop codons are indicated by asterisks.

Fig. 3. Schematic representation of the expression vectors pURI2–Ery (7659 bp) (A) and pURI3–Ery (2564 bp) (B). The expression leader sequences are the sequences present on pURI2 and pURI3 vectors, respectively [5].

Fig. 4. (A) Schematic representation of the expression vector pURI2–Cter (7755 bp). (B) Sequence of the pURI2–Cter sequence showing the C–terminus His_6–tag followed by four stop codons arranged in tandem. The expression leader sequence is under the control of the lpp^p–5 and lac^{PO} promoters. Restriction sites are indicated and underlined in the nucleotide sequence: X, *Xba*I, N, *Not*I, H, *Hin*dIII, and B, *Bam*HI. Tandem stop codons are indicated by asterisks.

Fig. 5. (A) Schematic representation of the expression vector pURI3–Cter (2642 bp). (B) Sequence of the pURI3–Cter sequence showing the C–terminus His_6–tag followed by four stop codons arranged in tandem. The expression leader sequence is under the control of the T7 polymerase Ø promoter in pURI3–Cter. Restriction sites are indicated and underlined in the nucleotide sequence: Nd, *Nde*I, N, *Not*I, H, *Hin*dIII, and C, *Cla*I. Tandem stop codons are indicated by asterisks.

Fig. 6. Expression of the *pad* gene encoding the PAD protein from *L. plantarum* cloned into the new pURI expression vectors. SDS–PAGE analysis of soluble (A) and insoluble (B) cell extract fractions from IPTG–induced cultures, respectively. Lane 1, *E. coli* **DH5α** (pURI2); lane 2, *E. coli* **DH5α (pURI2**–PAD); lane 3, *E. coli* **DH5α (pURI2**–TEV–PAD); lane 4, *E. coli* **DH5α (pURI2**–Ery–PAD); lane 5, *E. coli* **DH5α (pURI2**–Cter–PAD); lane 6, *E. coli* JM109 (DE3) (pURI3); lane 7, *E. coli* JM109 (DE3) (pURI3–PAD); lane 8, *E. coli* JM109 (DE3) (pURI3–TEV–PAD); lane 9, *E. coli* JM109 (DE3) (pURI3–Ery–PAD); lane 10, *E. coli* JM109 (DE3) (pURI3–Cter–PAD). The positions of molecular mass markers (Bio–Rad) are indicated on the left. The position of the PAD protein is indicated by an arrow. The expected sizes of the fused PAD proteins are 22.9 kDa when produced in pURI– and pURI–Ery vector, 23.0 kDa when produced in pURI–TEV vectors, and 21.8 kDa when produced in pURI–Cter vectors.

Fig. 7. Relative proportion of soluble (A) or insoluble (B) *L. plantarum* PAD protein produced by the different pURI expression vectors: pURI2–PAD (2), pURI2–TEV–PAD (3), pURI2–Ery–PAD (4), pURI2–Cter–PAD (5), pURI3–PAD (7), pURI3–TEV–PAD (8),

pURI3–Ery–PAD (9), pURI3–Cter–PAD (10). The PAD protein was quantified from the SDS–PAGE gels by using the ChemiDoc XRS+ Imagin System (Bio–Rad). pURI2 (1) and pURI3 (6) vectors were used as controls for the quantification.

Fig. 8. Cleavage of the *L. plantarum* PAD protein by TEV protease. PAD was purified from pURI3–TEV–PAD extracts on a His–Trapp–FF crude chelating affinity column, eluted with buffer containing 70 mM imidazole, dialyzed overnight into 20 mM Tris–HCl, pH 8.0 buffer free of imidazole, and incubated with TEV protease at 4 ºC. Lane 1, untreated PAD protein (23.0 kDa); and lane 2, PAD protein after 20 h incubation (21.0 kDa). Broad–range molecular weight markers (Bio–Rad) are run on the left and some positions are indicated.

Fig. 9. Production of *L. plantarum* PAD into ten different colonies of E.coli JM109 (DE3) harbouring pURI2–TEV–PAD plasmid. SDS–PAGE analysis of soluble cell extracts from IPTG–induced cultures. Lane 1 to 10, colony 1 to 10.The positions of molecular mass markers (Bio–Rad) are indicated on the left. The position of the PAD protein is indicated by an arrow.

Fig. 10. Expression of the *pad* gene encoding the PAD protein from *L. plantarum* cloned into the pURI expression vectors by using *E. coli* JM109 (DE3) cells. SDS–PAGE analysis of soluble cell extracts from IPTG–induced cultures. Lane 1, pURI2; lane 2, pURI2–PAD; lane 3, pURI2–TEV–PAD; lane 4, pURI2–Ery–PAD; lane 5, pURI2–Cter–PAD; lane 6, pURI3; lane 7, pURI3–PAD; lane 8, pURI3–TEV–PAD; lane 9, pURI3–Ery–PAD; lane 10, pURI3–Cter–PAD. The positions of molecular mass markers (Bio–Rad) are indicated on the left. The position of the PAD protein is indicated by an arrow.

Table 1

Table 1

Table 1. Relevant properties of pURI vector family

Vector	Size (kb)	*E .coli* host	Promoter	His Tag	Protease site	Antibiotic resistance
pURI2	7.8	**DH5α, DH10B**	*lpp*^{P}–5	N–ter	EK	Amp
pURI2–TEV	7.8	**DH5α, DH10B**	*lpp*^{P}–5	N–ter	TEV	Amp
pURI2–Ery	7.7	**DH5α, DH10B**	*lpp*^{P}–5	N–ter	EK	Ery
pURI2–Cter	7.8	**DH5α, DH10B**	*lpp*^{P}–5	C–ter	EK	Amp
pURI3	2.7	BL21(DE3), JM109 (DE3)	Ø10	N–ter	EK	Amp
pURI3–TEV	2.7	BL21(DE3), JM109 (DE3)	Ø10	N–ter	TEV	Amp
pURI3–Ery	2.6	BL21(DE3), JM109 (DE3)	Ø10	N–ter	EK	Ery
pURI3–Cter	2.7	BL21(DE3), JM109 (DE3)	Ø10	C–ter	EK	Amp

Table 2

Table 2

Table 2. Oligonucleotides used for cloning into the pURI vector family

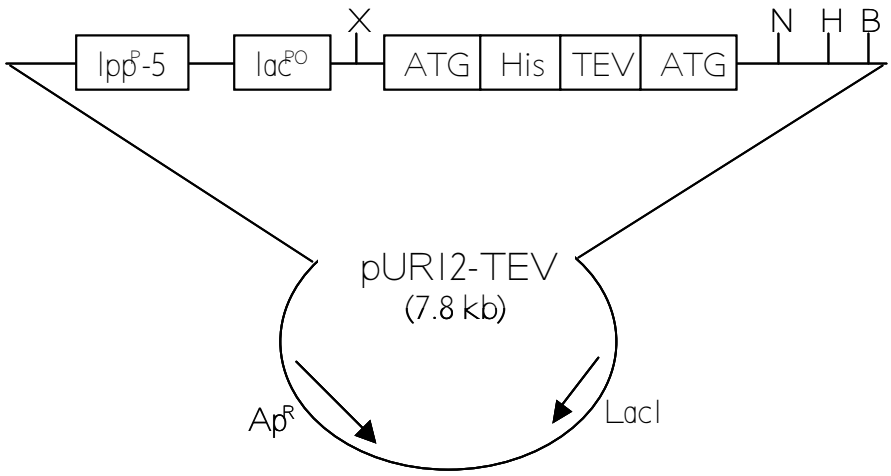
Oligonucleotide (5´–3´)[a]	pURI2	pURI2–TEV	pURI2–Ery	pURI2–Cter	pURI3	pURI3–TEV	pURI3–Ery	pURI3–Cter
*CATCAT*GGTGACGATGACGATAAGatg––	forward		forward		forward		forward	
GGTGAAAACCTGTATTTCCAGGGCatg––		forward				forward		
CTGGAACTCTAGAGGGTATTAATAatg––				forward				
TAACTTTAAGAAGGAGATATACATatg––								forward
AAGCTTAGTTAGCTATTATGCGTAtta––	reverse	reverse	reverse		reverse	reverse	reverse	
GCTATTA*ATGATGATGATGATGATG*––				reverse				reverse

[a] Only the sequence pairing the vector is indicated in the oligonucleotide. The start and stop condons are indicated in bold. The sequence coding for the poly–His tag is written in italics.

Figure 1

Figure 1

A



B

Ssp*I* lpp *promoter*

AATA**TTGACA**ACATAAAAAACTTTG**TG**T**TATACT**TGTAACGCTACATGGAGATTAACTCAATCTAGCTAG

-35 -10

***lac* promoter-operator**

AGAGGC**TTTACA**CTTTATGCTTCCGGCTCG**TATAAT**GTGTGGAATTGTGAGCGGATAACAATTTCACACA

-35 -10

*Xba*I **RBS**

GGAAACAGCTATGACCATGATTACGGATTCACTGGAACTCT**AGAGGG**TATTAATAATGGGGGGTTCT*CAT*

M G G S H

His tag **TEV recognition site**

*CATCATCATCATCAT*GGTGAAAACCTGTATTTCCAGGGC**ATG**CCTGCTACTGCTAATCGCTATCATTTTG

H H H H H G E N L Y F Q G M

GCGGGACACCTGGCATCAGCCCCGCTAACCCTGCGGCCCAAAATCGGGTATACTGAACCTAACTCTTAGT

AGAAAGTAGGTCAGCGTCATTGGCAAATTACATCAAGGAAATTCGCGAACTCGTCGGTCACAAACCCATC

*Not*I

ATTCTTAACGCTTCTGGTGGTCTGGTGACCAACGAGCGACACGAGGCTTTACTGAATCTGCGGCCGCTAC

*Hin*dIII *Bam*HI

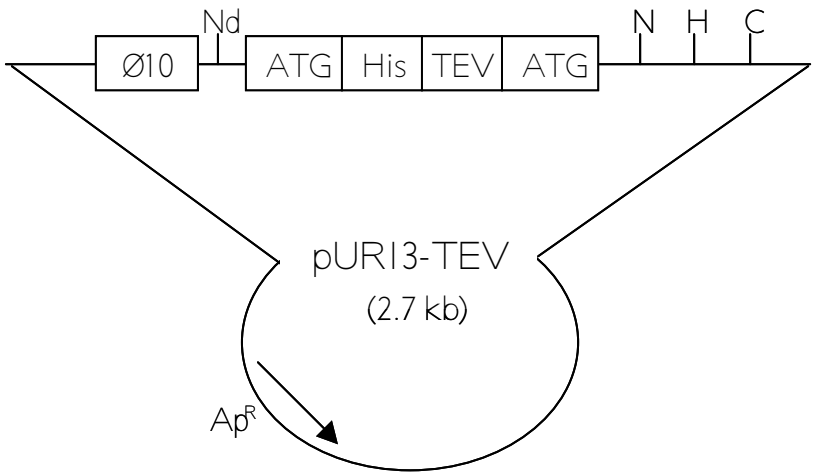
GCATAATAGCTAACTAAGCTTGGATCCGGCTGAGCAACGACGTGAACGCAA

* * * *

Figure 2

Figure 2

A



B

T7 promoter

GATAGACTTCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGT

RBS *Nde*I **His tag** **TEV recognition**

TTAACTTTAAG**AAGGAG**ATATACATATGGGGGGTTCT*CATCATCATCATCATCATCAT*GGTGAAAACCTGTAT

M G G S H H H H H H G E N L Y

site

TTCCAGGGC**ATG**CCTGCTACTGCTAATCGCTATCATTTTGGCGGGACACCTGGCATCAGCCCCGCTAACC

F Q G M

CTGCGGCCCAAAATCGGGTATACTGAACCTAACTCTTAGTAGAAAGTAGGTCAGCGTCATTGGCAAATTA

CATCAAGGAAATTCGCGAACTCGTCGGTCACAAACCCATCATTCTTAACGCTTCTGGTGGTCTGGTGACC

*Not*I *Hin*dIII *Cla*I

AACGAGCGACACGAGGCTTTACTGAATCTGCGGCCGCTACGCATAATAGCTAACTAAGCTTATCGATGAT

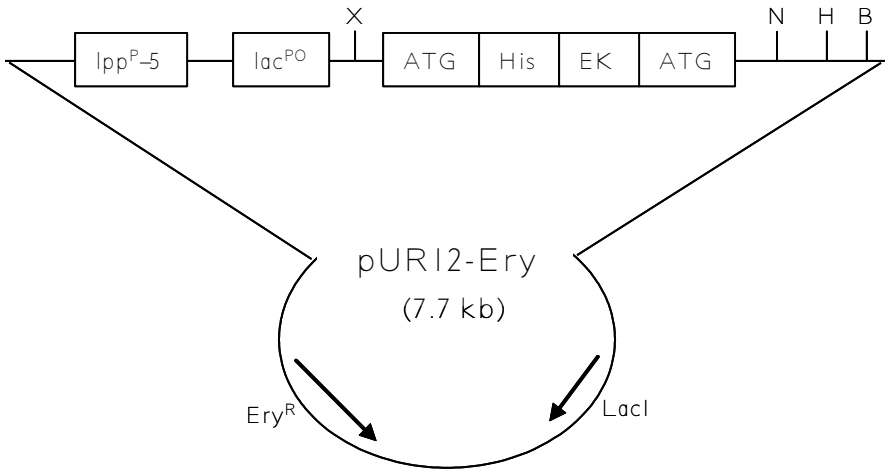
* * * *

AAGCTGTCAAACATGAGAATT

Figure 3

Figure 3

A



B

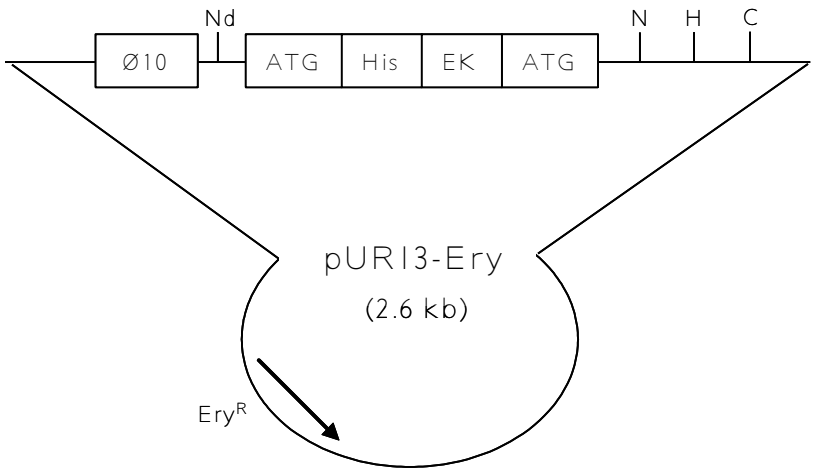
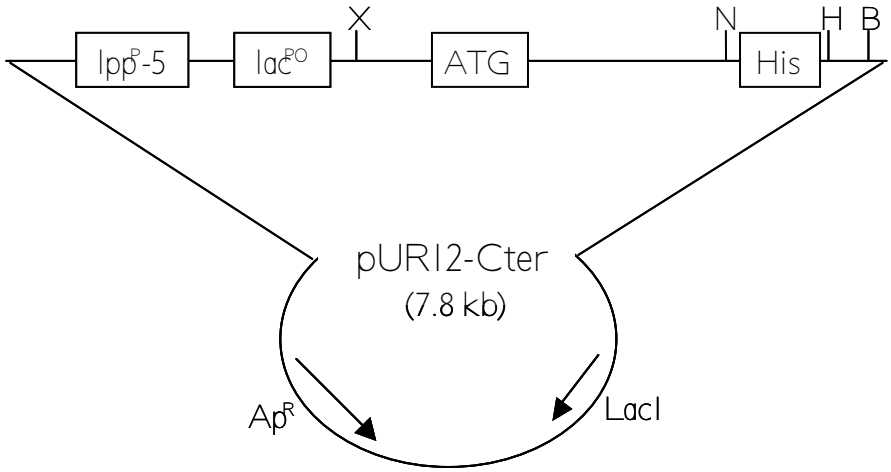


Figure 4

Figure 4

A



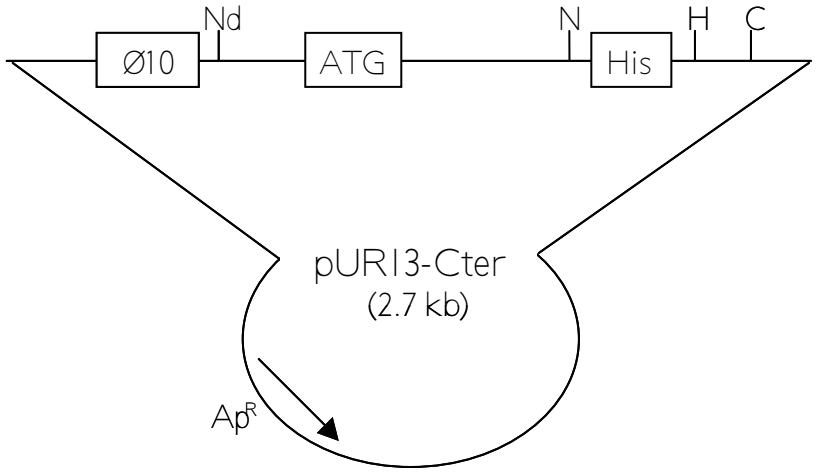
B

```
  SspI          lpp promoter
AATATTGACAACATAAAAAACTTTGTGTTATACTTGTAACGCTACATGGAGATTAACTCAATCTAGCTAG
      -35                     -10

                              lac promoter-operator
AGAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACA
       -35                     -10

                                        XbaI RBS
GGAAACAGCTATGACCATGATTACGGATTCACTGGAACTCTAGAGGGTATTAATAATGCCTGCTACTGCT
                                                 M

AATCGCTATCATTTTGGCGGGACACCTGGCATCAGCCCCGCTAACCCTGCGGCCCAAAATCGGGTATACT

GAACCTAACTCTTAGTAGAAAGTAGGTCAGCGTCATTGGCAAATTACATCAAGGAAATTCGCGAACTCGT

CGGTCACAAACCCATCATTCTTAACGCTTCTGGTGGTCTGGTGACCAACGAGCGACACGAGGCTTTACTG

       NotI          His tag                  HindIII  BamHI
AATCTGCGGCCGCTACGCACATCATCATCATCATCATTAATAGCTAACTAAGCTTGGATCCGGCTGAGCA
                  H  H  H  H  H  H  *  *   *   *

ACGACGTGAACGCAA
```

Figure 5

Figure 5

A



B

T7 promoter

GATAGACTTCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGT

RBS *Nde*I

TTAACTTTAAG**AAGGAG**ATATACAT**ATG**CCTGCTACTGCTAATCGCTATCATTTTGGCGGGACACCTGGC

M

ATCAGCCCCGCTAACCCTGCGGCCCAAAATCGGGTATACTGAACCTAACTCTTAGTAGAAAGTAGGTCAG

CGTCATTGGCAAATTACATCAAGGAAATTCGCGAACTCGTCGGTCACAAACCCATCATTCTTAACGCTTC

*Not*I **His tag**

TGGTGGTCTGGTGACCAACGAGCGACACGAGGCTTTACTGAATCTGCGGCCGCTACGCA*CATCATCATCA*

H H H H

*Hind*III *Cla*I

*TCATCAT*TAATAGCTAACTAAGCTTATCGATGATAAGCTGTCAAACATGAGAATT

H H * * * *

Figure 6

Figure 6

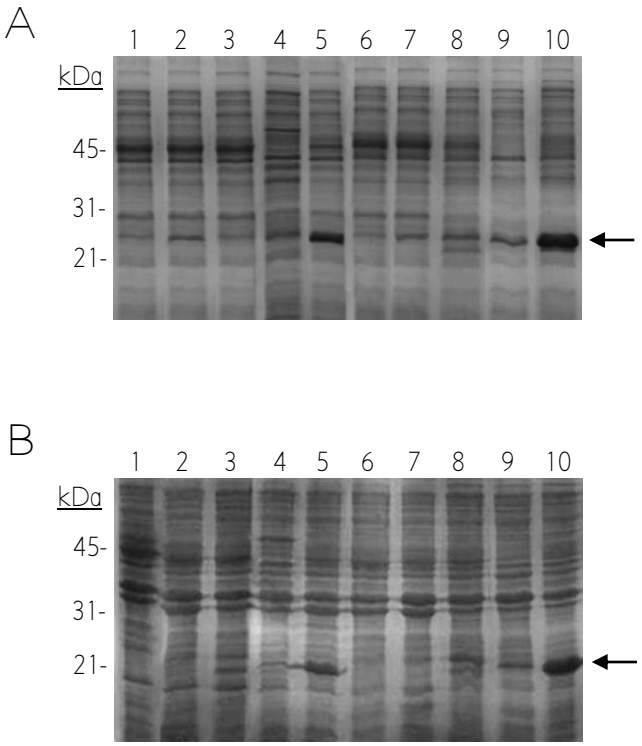


Figure 7

Figure 7

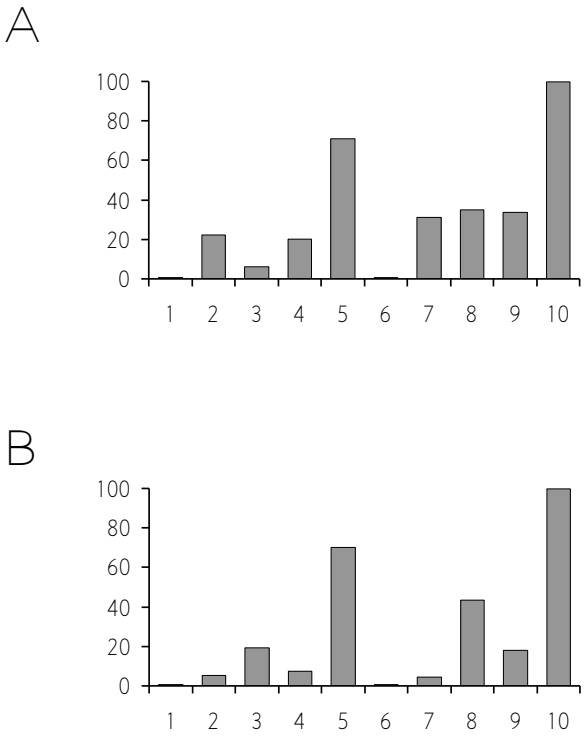


Figure 8

Figure 8

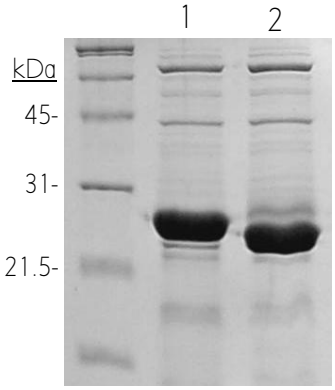


Figure 9

Figure 9

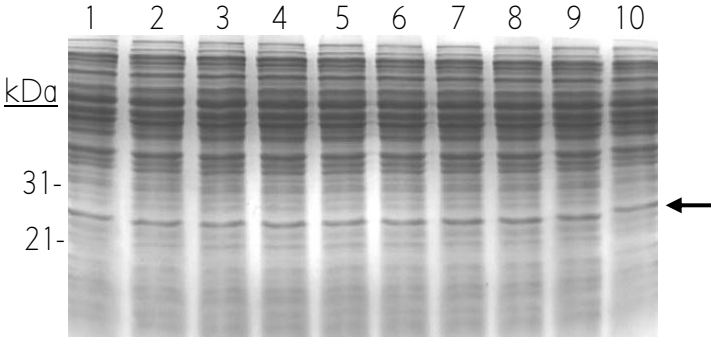


Figure 10

Figure 10