Low-copy-number T7 vectors for selective gene expression and efficient protein overproduction in *Escherichia coli*

Petra Dersch, Hafida Fsihi and Erhard Bremer *

Max-Planck-Institut für Terrestische Mikrobiologie, Karl-von-Frisch Str, D-35043 Marburg, FRG

(Received 15 July 1994, revision received 29 July 1994, accepted 4 August 1994)

Abstract A set of low-copy-number vectors (pPD) has been constructed that permit selective gene expression and high-level protein overproduction in *Escherichia coli*, based on the bacteriophage T7 RNA polymerase/T7 promoter system. These plasmids carry a chloramphenicol resistance gene (*cat*) as a selective marker and an extended multiple cloning site for convenient gene cloning. Their replication is mediated by ori sequences derived from the low-copy-number vector pSC101. The efficient T7 gene 10 promoter present on these vectors allows selective and high-level transcription of cloned genes carrying their own translational initiation signals. In addition, low-copy-number T7 vectors were constructed that permit expression of genes lacking their own transcription and translation initiation elements by providing a ribosome binding site, an ATG start codon and a multiple cloning site devised for the cloning in all three reading frames. The pPD expression vectors were used to achieve high-level overproduction of the *E. coli* integral outer membrane protein Tss, and the cytoplasmic enzymes β-galactosidase (βGal) and UTP α-D-glucose-1-phosphate uridylyltransferase (GalU). The characteristics of these low-copy-number T7 expression vectors should prove very useful for the cloning and high-level overexpression of genes whose gene products are deleterious to the *E. coli* host.

Key words Phage T7 gene 10 promoter, T7 RNA polymerase, Selective gene expression, Protein overproduction, pSC101 replicon, *Escherichia coli*

Introduction

The T7 RNA polymerase/T7 promoter system established by Tabor and Richardson [1] and Studter and Moffatt [2] is an extremely valuable tool for high-level gene expression and efficient protein overproduction. In recent years, several versatile T7 expression vectors have been developed and refined which contain the strong promoter of the T7 gene 10 and allow the positioning of target genes under the transcriptional control elements of the T7 promoter [3]. Most of the described T7 expression vectors are plasmids with a high-copy-number [3], and this feature can greatly complicate the stable cloning of genes whose products are deleterious for cell growth and viability. We experienced such a situation during the characterization of the *Escherichia coli*
**Materials and Methods**

**Growth conditions and bacterial strains**

Bacteria were grown aerobically at 37°C either in rich LB medium or in M9 minimal medium with 0.2% glucose as the carbon source [6,7]. Plasmids were propagated in strain MC4100 [7]. T7 RNA polymerase mediated overproduction of proteins was done in strain BL21 (ADE3) [2]. Strain HF18 is a tsx Tn10 derivative of strain BL21 and was used for plasmid pHF1 (tsx') [8] mediated overproduction of the *E. coli* Tsx protein.

**DNA manipulations**

All manipulations of nucleic acids followed routine procedures [7]. The nucleotide sequence of the multiple cloning sites of plasmids pPD1,
pPD100, pPD101, pPD110, pPD111 and pPD112 (Fig 1) were confirmed by double-stranded DNA sequencing with the chain-termination method [9] using the lacZ (−40) primer.

**Plasmid constructions**

For the construction of pPD1 (Fig 1A), an 89-bp *PvuII-HindIII* fragment of the plasmid pT7-5 [1] containing the T7 gene 10 promoter region was isolated and ligated into the *SmaI* and *HindIII* site of plasmid pHSG575 [10]. For the construction of plasmids pPD100 and pPD101, the polylinker sequences of the T7 expression vectors pT7-5 and pT7-6 [1] were exchanged with an 87-bp *EcoRI-HindIII* DNA fragment carrying the extended multiple cloning site of plasmid pUC-BM20 (Boehringer Mannheim), resulting in the vectors pPD98 and pPD99. In a second step, a *PvuII-HindIII* fragment (134 bp) was isolated from pPD98 and inserted into the *MunI* and *HindIII* sites of pHSG575 [10], yielding plasmid pPD100 (Fig 1B). To reverse the orientation of the multiple cloning site, a *PvuII-EcoRI* DNA fragment (133 bp) was prepared from pPD99 and cloned into the *MunI* and *EcoRI* sites of pHSG576 [10], yielding plasmid pPD101 (Fig 1B).

For the construction of T7 expression vectors providing translation initiation signals, we amplified by polymerase chain reaction the translation initiation elements from plasmids pTrc97A, pTrc97B and pTrc97C [11]. This DNA segment was amplified by use of primer 1 (5'-GTG AGC GGA TAT CAA TTT CAC ACA GG-3') which generates an *EcoRV* restriction site, and primer 2 (5'-GCG GTG AGC TCC TTC GGT CC-3') which generates an *EcoRV* restriction site, and primer 2 (5'-GCG GTG AGC TCC TTC GGT CC-3'). The obtained PCR fragments were digested with *EcoRV* and *HindIII* and cloned into pPD100 (Fig 1B). These constructions yielded plasmids pPD110 (3 525 kb), pPD111 (3 526 kb) and pPD112 (3 527 kb) (Fig 1C).

A lacZ gene lacking its promoter was isolated as a 4 35-kb *DraI* restriction fragment from pMLB1010 [7] and inserted into the *EcoRV* site of plasmids pPD100 and pPD101 (Fig 1B), generating pPD129 and pPD130, respectively. To construct a plasmid expressing a lacZ gene lacking its own transcriptional and translational initiation signals, a 3 14-kb *EcoRI-DraI* fragment from pNM482 [12] was isolated and ligated into vector pPD112 (Fig 1C), which had been digested with *EcoRI* and *SmaI* yielding pPD124. A 7 90-kb *PstI* fragment encoding the *galU* gene was isolated from pGM7 [4,13] and inserted into the unique *PstI* site of pPD110 (Fig 1B) yielding plasmid pPD114.

**T7 promoter mediated protein overproduction**

For the overproduction of the *E. coli* outer membrane protein Tsx, fresh overnight cultures of strain HF18 (pHF1), were diluted 100-fold either into M9 minimal medium with 0.2% glucose as the carbon source or into LB-rich medium with chloramphenicol (30 μg ml⁻¹) to select for the pHF1 plasmid. The cultures were grown at 37°C in a shaking water bath to mid-log phase (*A₅₇₈ = 0.5–1*), and the expression of the chromosomal T7 RNA polymerase gene under lacP₀ control was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After 30 min, rifampicin (200 μg ml⁻¹) was added to the cultures to inhibit the *E. coli* RNA polymerase [3], and the cultures were shaken for an additional hour. Aliquots (1 ml) were removed before and after induction of the expression of the T7 RNA polymerase gene from each culture and the cells were pelleted by centrifugation. From the M9- and LB-grown cells a crude outer membrane preparation was prepared from 35 ml culture [14]. Proteins from both the total cell extracts and the cell envelope proteins were resuspended in 40 μl SDS sample buffer, solubilized by boiling for 5 min, and loaded onto a 12% SDS polyacrylamide gel. The cellular proteins were visualized by staining with Coomassie blue and Tsx-related polypeptides were immunologically identified after transfer on an Immobilon membrane with a Tsx antiserum [14]. For the overproduction of the cytoplasmic enzymes β-galactosidase (βGal) and UTP α-D-glucose-1-phosphate undylyltransferase (GalU), plasmids carrying the corresponding lacZ and *galU* genes were transformed into strain BL21 (ADE3) and used for T7-mediated gene expression in cells grown in M9 minimal medium with 0.2% glucose as the carbon source. The cultures were grown to mid-log phase and βGal and GalU overprodu-
tion was performed as described above for the Tsx protein. Aliquots (1 ml) were removed before and after induction from each culture. Whole cellular extracts were prepared and resuspended in 40 μl sample buffer [7]. The protein extracts of the βGal overproducing cultures were electrophoretically separated on 7% SDS polyacrylamide gels and those from GalU synthesizing cells were loaded on a 12% SDS polyacrylamide gel. Proteins were visualized by staining with Coomassie blue.

Results and Discussion

Construction of the T7 expression vector pPD1 and overexpression of the Tsx protein

For the construction of low-copy-number vectors carrying the strong gene 10 promoter of phage T7, we used plasmid pHSG575 as the vector backbone. This small plasmid (3.61 kb) confers resistance to chloramphenicol and contains a multiple cloning site within the lacZ α gene derived from pUC8 [10]. Replication of pHSG575 is mediated by the pSC101 ori sequences, yielding 6 to 8 copies of the plasmid per cell [15]. A small restriction fragment carrying the strong phage T7 gene 10 promoter was isolated from the multi-copy plasmid pT7-5 [1] and inserted into pHSG575, resulting in the low-copy-number expression vector pPD1 (Fig 1A).

To test the utility of our T7 expression system for overproducing heterologous proteins, we used the E. coli tsx gene as an example. The tsx gene encodes an integral outer membrane protein (Tsx, 272 aa, M, 31418) that functions as a substrate-specific channel for nucleosides and deoxynucleosides and is initially synthesized with a 22-aa bacterial signal sequence peptide (Pro-Tsx) [8,14]. Like other integral proteins from the E. coli outer membrane, Tsx is deleterious to cell growth when overproduced [14]. Plasmid pHF1 carries the tsx gene, including its own transcriptional and translational regulatory sequences, it therefore mediates low-level tsx expression even under non-inducing conditions [8,14]. For the overexpression experiments, plasmid pHF1 was transformed into strain HF18, a tsx Tn10 derivative of strain BL21 (λDE3), carrying a chromosomal copy of the T7 RNA polymerase structural gene (T7 gene 1) under the control of the lac promoter/operator regulatory elements [2]. Induction of the T7 promoter-directed tsx expression resulted in the appearance of a predominant polypeptide species in whole cell extracts and in the crude cell envelope preparations (Fig 2A) whose identity with the Tsx protein was proven by Western blot analysis (Fig 2B). Plasmid pHF1-directed overexpression of Tsx also caused the massive accumulation of the Pro-Tsx protein which can normally not be detected due to its efficient processing. The high rate of Pro-Tsx production overwhelms the cellular machinery that mediates Tsx secre-
tion and thus causes the pronounced accumulation of the Pro-Tsx protein.

Construction and properties of the low-copy-number T7 expression vectors pPD100 and pPD101

The data presented above show that the low-copy-number T7 promoter plasmid pPD1 can be used for the efficient overproduction of proteins in E. coli. To improve the utility of this expression vector system, we expanded the multiple cloning site and removed the lac promoter present in pPD1 (Fig 1A). These plasmid constructions yielded two chloramphenicol resistant low-copy-number T7 expression vectors, pPD100 (3.52 kb) and pPD101 (3.52 kb), allowing convenient cloning of heterologous genes in both orientations (Fig 1B). To test these vectors, we inserted a promoterless lacZ gene into pPD100 and pPD101 and expressed it in strain BL21 (λDE3) under T7 control. Both plasmids directed the massive synthesis of β-galactosidase (βGal) under inducing conditions, but there was very little βGal production in non-induced cells (Fig 3A). Thus, both pPD100 and pPD101 can be used to achieve strong overproduction of recombinant proteins pPD110, pPD111 and pPD112. T7 expression plasmids for placing cloned target genes under vector provided translational initiation signals

Using pPD100 (Fig 1B) as the plasmid backbone we developed a set of low-copy-number T7 vectors that provide translation initiation elements that are followed by a multiple cloning site and permit subcloning in any one of the three reading frames. For the construction of these plasmids, we capitalized on the multi-copy pTrc97A, pTrc97B and pTrc97C expression vectors that carry the lacZ ribosome binding site, an appropriately positioned ATG start codon and a modified multiple cloning site of pUC18 [11]. DNA segments carrying the translational initiation elements and the adjacent multiple cloning site of pTrc97A, pTrc97B, and pTrc97C were amplified by polymerase chain reaction and inserted into the multiple cloning site of pPD100, yielding the low-copy-number expression vectors pPD110, pPD111 and pPD112 (Fig 1C). To study

![Fig 3 Overexpression of the βGal and GalU proteins](image)

Lanes 1, 3 and 5 show whole protein extracts from non-induced and lanes 2, 4 and 6 from induced cell cultures (A) Lanes 1, 2 BL21 (pPD100, vector), lanes 3, 4 BL21 (pPD129, a derivative of plasmid pPD100 carrying a 'lacZ gene expressed from its own translation initiation signals), lanes 5, 6 BL21 (pPD130, a derivative of plasmid pPD101 carrying a 'lacZ gene expressed from its own translation initiation signals) (B) Lanes 1, 2 BL21 (pPD124, a derivative of plasmid pPD110 carrying a 'lacZ gene expressed from the translation initiation signals provided by the T7 expression vector) (C) Lanes 1, 2 BL21 (pPD110, vector), lanes 3, 4 BL21 (pPD114, a derivative of plasmid pPD110 carrying the galU gene expressed from its own translation initiation signals) The arrows point to the overproduced βGal and GalU proteins.
the level of T7-mediated gene expression by these vectors, we first cloned the *E. coli* galU gene carrying its own transcription and translation initiation signals into plasmids pPD110, pPD111 and pPD112. The galU gene encodes a cytoplasmic UTP α-β-glucose-1-phosphate uridylyltransferase (GalU) [13]. Following the induction of the T7 expression system, the GalU protein was very strongly produced, an example for this overproduction is documented in Fig 3C. Thus, plasmids pPD110, pPD111 and pPD112 can also be used as efficient transcription T7 expression vectors, similar to low-copy-number T7 promoter plasmids characterized above. To test the use of these plasmids for the expression of genes lacking their own transcription and translation initiation elements, we fused a *lacZ* gene in all three reading frames to the ATG start codon present in the multiple cloning site of plasmids pPD110, pPD111 and pPD112 (Fig 1C) and monitored βGal synthesis in strain BL21 (ADE3) βGal was readily detected in whole cell extracts following induction of T7-mediated lacZ expression (Fig 3B)

### Conclusions

The low-copy-number T7 promoter vectors described here have a number of desirable features that make them interesting alternatives to other T7 expression vectors [1–3].

(1) Their low-copy-number and the extended multiple cloning site permit the cloning of genes whose gene products are deleterious to the *E. coli* host and which therefore can often not be stably established in high-copy-number expression plasmids. The utility of plasmid pPD1 for the overproduction of toxic proteins has already been proven. We have used this plasmid successfully to massively overproduce the *E. coli* nucleoid-associated DNA binding protein H-NS [5], which is highly detrimental for cell growth and viability even when expressed from its own promoter on a low-copy-number plasmid [4]. Others, have employed pPD1 to clone and express *urK* an essential virulence gene for the intracellular spreading of *Shigella flexneri*, which could not be cloned into the multi-copy T7 expression vector pT7–5 [18].

(2) Although replication of the pPD-T7 vectors is mediated by ori sequences derived from the low-copy-number pSC101 replicon, high-level heterologous gene expression under T7 control can be reliably achieved. We observed low-level basal gene expression in non-induced cell cultures and efficient protein overproduction after induction of T7 RNA polymerase synthesis. Several special-purpose low-copy-number plasmids have been constructed which contain the T7 gene 10 promoter [16,17]. However, the utility of low-copy number vectors for high-level overproduction of recombinant proteins has not been widely appreciated.

(3) The pSC101 ori of the pPD-T7 vectors is compatible with the frequently used high-copy-number CoIE1-derived plasmids and can thus co-exist with other vectors in the same cell allowing co-overproduction of different proteins. Commercially available M13 *lacZ α* and T7 primers permit double-stranded DNA sequencing of the cloned DNA fragments.

### Acknowledgements

We thank A Weissborn and E Kennedy for communicating data prior to publication. The expert technical assistance of S Kneip and the help of V Koogle in preparing the manuscript are greatly appreciated. We thank R Thauer for his continued support. Financial support for this study was provided by the Max-Planck Gesellschaft and the Fonds der Chemischen Industrie. P D and H F gratefully acknowledge the receipt of fellowships from the Boehringer Ingelheim Foundation and the Deutsche Akademische Austauschdienst, respectively.

### References