# Exploitation of the Escherichia coli lac operon promoter for controlled recombinant protein production

Browning, Douglas F; Godfrey, Rita E; Richards, Kirsty L; Robinson, Colin; Busby, Stephen J W

*DOI:*
10.1042/BST20190059



*Document Version*
Peer reviewed version

*Citation for published version (Harvard):*
Browning, DF, Godfrey, RE, Richards, KL, Robinson, C & Busby, SJW 2019, 'Exploitation of the Escherichia coli lac operon promoter for controlled recombinant protein production', *Biochemical Society Transactions*, vol. 47, no. 2, pp. 755-763. https://doi.org/10.1042/BST20190059

For Biochemical Society Transactions **Edited by DFB (11/03/2019)**

**BST-2019-0059C**

# Exploitation of the *Escherichia coli lac*-operon promoter for controlled recombinant protein production

Douglas F. Browning[1], Rita E. Godfrey[1], Kirsty L. Richards[2], Colin Robinson[2] and Stephen J. W. Busby[1]

[1] Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Birmingham, B15 2TT, UK.
[2] School of Biosciences, University of Kent, Ingram Building, Canterbury CT2 7NJ, UK

**Correspondence:** Douglas F. Browning (D.F.Browning@bham.ac.uk)
Stephen J. W. Busby (S.J.W.Busby@bham.ac.uk)

**Word Count: 2183 in main text**

**Abstract**

The *Escherichia coli lac* operon promoter is widely used as a tool to control recombinant protein production in bacteria. Here we give a brief review of how it functions, how it is regulated, and how, based on this knowledge, a suite of *lac* promoter derivatives has been developed to give controlled expression that is suitable for diverse biotechnology applications.

**Key words:** *Escherichia coli*, bacterial transcription initiation, *lac* promoter, *tac* promoter, Lac repressor, recombinant protein production

**Abbreviations Used:** RPP, recombinant protein production; RNAP, RNA polymerase; bp, base pairs; CRE, Core-recognition element; CRP, cyclic AMP receptor protein; LacI, lactose operon repressor protein; IPTG, isopropyl β-D-1-thiogalactopyranoside; PAR, promoter activity rating; GFP, green fluorescent protein; hGH, human growth hormone; Tat, twin arginine translocon pathway; LB: lysogeny broth; ONPG, o-nitrophenyl-β-D-galactopyranose.

## Introduction

Most of the recombinant protein production (RPP) systems used for expressing proteins in bacteria were constructed in the last century [1-4]. High level RPP provided by these systems enables the synthesis and purification of large amounts of soluble recombinant protein. However, the expression of difficult protein targets (*e.g.* membrane proteins or proteins secreted out of the cytoplasm), using these RPP systems, may be too high for cells to cope and adequately fold protein, resulting in substantial target degradation or the production of insoluble aggregates (*i.e.* inclusion bodies) [5-8]. Many "tricks of the trade" can be employed to slow down RPP expression and increase the level of soluble product, *e.g.* lowering the growth temperature, decreasing the inducer concentration or using a weaker promoter [6-9]. Although such tinkering can be very successful, determining the correct combination of refinements can be time-consuming. This can also be very "hit-and-miss", being dependent on the particular target protein in question [6, 9], and for some induction regimes, for example when using low inducer concentrations, only a proportion of the cells in a culture may in fact express recombinant protein [10-12].

The *E. coli lac* operon promoter was one of the first bacterial promoters to be adopted by biotechnologists for RPP, and it is still used today, especially when *Escherichia coli* is used as the host [6]. Here, we give a brief update of our current understanding of transcript initiation in bacteria, emphasising special features of the *lac* promoter and its regulation. We then review how, based on this information, many *lac* promoter derivatives have now been engineered in order to facilitate controlled RPP expression and avoid the problems that are concomitant with high level overexpression.

## Transcript initiation and regulation at the *E. coli lac* operon promoter

Transcript initiation in bacteria takes place when the multisubunit DNA-dependent RNA polymerase holoenzyme (RNAP) interacts with a DNA promoter sequence (Figure 1A). In brief, the RNAP first interacts with double-stranded DNA to form a 'closed complex' in which determinants in different RNAP subunits interact with different promoter sequence elements (Figure 1B) [13]. Thus a determinant in the RNAP α subunit C-terminal domain interacts with the promoter UP element, a determinant in Domain 4 of the RNAP σ subunit interacts with the promoter -35 element, and a determinant in Domain 3 of the RNAP σ subunit interacts with the extended -10 element (Figure 1B). Following this, Domain 2 of the RNAP σ subunit drives the local unwinding of 13-15 base pairs (bp) of promoter DNA to form the 'open complex', in which the single-stranded DNA template strand is positioned in the active site of the RNAP, such that initiation of DNA-templated RNA synthesis can take place (Figure 1C) [14]. Formation of the initiation-competent 'open complex' from the 'closed complex' is driven by further specific interactions between other RNAP determinants and promoter sequences

[15-17]. Thus, determinants in Domain 2 of the RNAP σ subunit interact with single-stranded bases of the promoter -10 element and with the promoter discriminator element. These interactions involve only the non-template strand of the locally unwound segment of promoter DNA, thereby permitting the single-stranded template strand to access the RNAP active site. The exact position of the template strand in the active site, and the location of the downstream junction between single-stranded and double stranded DNA, is set by other contacts involving amino-acid side-chains of the RNAP β and β' subunits (which interact with the promoter Core-recognition element, CRE) (Figure 1C) [16].

The activity of any bacterial promoter is set by the formation of the 'closed complex' and the ensuing isomerisation to, and escape from, the 'open complex', as the RNAP copies the template strand to elongate its transcript. In the case of the *E. coli lac* operon promoter, defects in the UP element, -35 element and extended -10 element hinder 'closed complex' formation but this is remedied by binding of an activatory factor, the cyclic AMP receptor protein (CRP) to a target sequence centred between bp 61 and 62 (referred to as position -61.5) upstream of the transcript start (denoted as +1; see Figure 1D) [18, 19]. A second regulator protein, the lactose operon repressor protein (LacI) binds to a high affinity target sequence, known as operator O1, centred at position +11 (Figure 1D) [19, 20]. LacI binding to its target effectively shuts down *lac* promoter activity, but repression can be broken by the presence of allolactose (a breakdown product of lactose) or by the sugar analogue IPTG (isopropyl β-D-1-thiogalactopyranoside), which both bind to the LacI repressor and cause it to release operator DNA [20]. LacI-dependent repression of the *lac* promoter is supported by LacI binding to two secondary weaker operators, O3, located at position -82 (Figure 1D), and far downstream, O2 [19-23]. Here, we review how the starting *lac* promoter has been engineered to make it fit for different biotechnology purposes. We focus on constructions that release the requirement for CRP, base changes in key promoter elements, and the exploitation of the O3 operator.

**Activator-independent *lac* promoter derivatives**

Figure 2A illustrates the organisation of the *E. coli lac* operon promoter, showing key promoter elements, the DNA site for CRP, and the location of LacI-binding operators, O1 and O3. Figure 2B shows the base sequence of a typical DNA fragment carrying the *lac* promoter that might be used in any biotechnology application. The fragment, which carries a useful restriction site at each end, is denoted *lac* O3O1. Potentially, regulation by CRP might be exploitable but, since CRP activity is difficult to control by external cues, many biotechnology applications that use the *lac* promoter have sought to eliminate CRP effects and focussed on regulation by LacI. One way to do this is by the use of the *lac* UV5 mutant promoter. This mutant promoter carries a 2 bp change in the promoter -10 hexamer element that creates a consensus -10 promoter element (Figure 2B) [19, 24]. The alternative is to replace *lac* promoter upstream DNA sequences and this has been done in the *tac* promoter, which is a chimeric fusion between the upstream elements of the *E. coli trp* promoter and the downstream elements of the *lac*UV5 promoter [1, 2]. Figure 2A illustrates both the *lac*UV5 and *tac* promoters, and Figure 2B shows the base sequence of DNA fragments carrying these promoters. Note that the *tac* promoter carries consensus -35 and -10 promoters elements and a single operator for LacI, O1. For *E. coli* promoters that are dependent on the housekeeping sigma factor $\sigma^{70}$ the optimal spacing between the -35 and -10 promoter elements is 17 bp and deviation from this leads to a reduction in promoter activity [25, 26]. Thus, it worth noting that both the *tac* and *lac* promoters are in fact suboptimal, having spacers of 16 and 18 bp, respectively (Figure 2B).

Figure 3A-C illustrates the results of simple assays to compare the activities of the *tac*, *lac*UV5 and *lac* promoters. To perform these assays, the *tac* O1 fragment, the *lac*UV5 O3O1 fragment and the *lac* O3O1 fragment were each cloned into a plasmid expression vector (pRW50) such that the promoters controlled transcription of the gene (*lacZ)* encoding β-galactosidase [27]. Recombinant plasmids were then transformed into a *Δlac E. coli* host

strain and β-galactosidase expression was measured. As expected the hierarchy of promoter activity was *tac* > *lac*UV5 > *lac* (Figure 3).

**Modulation of promoter activities using upstream *lac* operator sequences**

A key feature of LacI-dependent repression of the *E. coli lac* promoter is the contribution of the auxiliary upstream operator, O3, and this has been exploited to tailor promoter activity levels. Data in Figure 3A illustrate how the introduction of certain *lac* operator sequences (*i.e.* O1 or the high affinity "ideal" *lac* operator Oid, see Figure 2A [21, 28, 29]) at position -82 of the *tac* O1 promoter fragment reduced the high activity of the *tac* promoter. Thus, expression from the starting *tac* O1 promoter fragment and each of the *tac* O3O1, *tac* O1O1 and *tac* OidO1 derivatives was induced by IPTG but the introduction of the O1 and Oid operator sequences decreased IPTG-induced expression levels. Similarly, data in Figures 3B and 3C illustrate how the introduction of higher affinity *lac* operators into the upstream region of the *lac*UV5 promoter (*i.e.* in the *lac*UV5 O1O1 and *lac*UV5 OidO1 promoter fragments) or the *lac* promoter (*i.e.* in the *lac* O1O1 and *lac* OidO1 promoter fragments) decreased promoter activity.

**The power of *lac*: combinations make anything possible**

In addition to controlling promoter activity by upstream-bound LacI, activity can be modulated by point mutations in different promoter elements. Data in Figure 3D illustrate how the p34G, p14G, p9A or p8A substitutions (at positions -34, -14, -9 or -8), which make the -35 element, the extended -10 or the -10 element more similar to the respective consensus, can be combined with different operator combinations to produce a suite of IPTG-inducible promoters with a wide range of activities.

To illustrate the use of these promoters in RPP, we selected a subset of 8 promoters from the above suite and gave each a promoter activity rating (PAR) value of PAR1 to PAR8 (Figure 4A). Our rationale for this is that, depending on the target protein being expressed, specific IPTG-induced expression levels could be achieved. To examine this, some of the promoter constructs were introduced into the low-copy-number vector pTorA-GFP [30] and the high-copy-number vector pHAK1 [31], using standard techniques [32-34]. For pTorA-GFP derivatives, each PAR promoter drives the expression of a *torA-GFP-6his* fusion (GFP, green fluorescent protein), whilst, for pHAK1, each construct expresses a *torA-hGH-6his* fusion (hGH, human growth hormone). Note that the *torA* signal sequence in each case will direct the recombinant protein to the Tat (twin arginine translocon) system for periplasmic targeting [7, 30, 31]. Plasmid constructs were transferred into *E. coli* BL21 cells and RPP was induced in bacterial cultures by the addition of 1 mM IPTG for 3 hrs [7]. Normalised total cellular protein from cells, expressing either TorA-GFP-6His or TorA-hGH-6His, were then analysed by Western blotting. Results in Figures 4B and 4C show that both GFP and hGH were expressed using the PAR promoter constructs. Note that, in some instances, two product bands can be observed, in each case the species with the higher molecular weight still carries the TorA signal sequence, whilst the smaller species has been processed and lacks the TorA moiety [7, 30, 31, 35]. Importantly, using this expression system, inducible expression of different target proteins can be set to specific levels when using both low- and high-copy-number vectors.

## Perspectives

The *lac* promoter and its derivatives have been widely used to express many recombinant proteins to high levels in *E. coli* [1, 2, 7, 36] and many currently used vectors have been designed to optimise expression. However, there are situations where expression must be moderated. For example, the secretion of recombinant biopharmaceuticals out of the *E. coli* cytoplasm into the periplasm is often a preferred industrial strategy, as this minimises downstream processing costs, since the target protein can be purified from the periplasmic contents, with minimal cellular and DNA contamination [37]. For this to be successful, RPP needs to be slowed down so that product is not degraded before it is transported [7].

In this review, we have sought to show how knowledge of the *E. coli lac* promoter can be exploited to produce a suite of derivative promoter fragments to cope with any situation. Previous promoter engineering had focused on altering the *lac* promoter -10 and -35 elements to change the basal promoter expression of constructs [1, 38]. Recent advances in our understanding of transcript initiation in bacteria and its regulation now allow *lac* promoters to be constructed with different operator sequences to alter the induced level of expression. This is possible because, even in the presence of inducer, LacI has some affinity for its operator sequence and, thus, in the induced state the LacI repressor can still remain bound to DNA, as has been observed in single molecule studies [39].

**Funding**

This work was generously supported by an Industrial Biotechnology Catalyst (Innovate UK, BBSRC, EPSRC) grant BB/M018261/1 to support the translation, development and commercialisation of innovative Industrial Biotechnology processes.

**Author contributions**

D.F.B., S.J.W.B. and C.R. conceived and designed the research programme. D.F.B, R.E.G and K.L.R. performed the experiments. D.F.B. wrote the manuscript with input from all authors.

**Competing Interests**

The Authors declare that there are no competing interests associated with the manuscript.

## References


1 de Boer, H.A., Comstock, L.J. and Vasser, M. (1983) The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. *Proc Natl Acad Sci U S A*. **80**, 21-25

2 Amann, E., Brosius, J. and Ptashne, M. (1983) Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene*. **25**, 167-178

3 Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol*. **177**, 4121-4130

4 Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*. **189**, 113-130

5 Wagner, S., Klepsch, M.M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K.J., Slotboom, D.J., Persson, J.O. and de Gier, J.W. (2008) Tuning *Escherichia coli* for membrane protein overexpression. *Proc Natl Acad Sci U S A*. **105**, 14371-14376

6 Overton, T.W. (2014) Recombinant protein production in bacterial hosts. *Drug Discov Today*. **19**, 590-601

7 Browning, D.F., Richards, K.L., Peswani, A.R., Roobol, J., Busby, S.J.W. and Robinson, C. (2017) *Escherichia coli* 'TatExpress' strains super-secrete human growth hormone into the bacterial periplasm by the Tat pathway. *Biotechnol Bioeng*. **114**, 2828-2836

8 Quick, M. and Wright, E.M. (2002) Employing *Escherichia coli* to functionally express, purify, and characterize a human transporter. *Proc Natl Acad Sci U S A*. **99**, 8597-8601

9 Gopal, G.J. and Kumar, A. (2013) Strategies for the production of recombinant protein in *Escherichia coli*. *Protein J*. **32**, 419-425

10 Novick, A. and Weiner, M. (1957) ENZYME INDUCTION AS AN ALL-OR-NONE PHENOMENON. *Proc Natl Acad Sci U S A*. **43**, 553-566

11 Maloney, P.C. and Rotman, B. (1973) Distribution of suboptimally induces -D-galactosidase in *Escherichia coli*. The enzyme content of individual cells. *J Mol Biol*. **73**, 77-91

12 Khlebnikov, A., Risa, O., Skaug, T., Carrier, T.A. and Keasling, J.D. (2000) Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. *J Bacteriol*. **182**, 7029-7034

13 Murakami, K.S. and Darst, S.A. (2003) Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol*. **13**, 31-39

14 Feklistov, A., Sharon, B.D., Darst, S.A. and Gross, C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu Rev Microbiol*. **68**, 357-376

15 Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R. and Darst, S.A. (2015) Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *Elife*. **4**

16 Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E. and Ebright, R.H. (2012) Structural basis of transcription initiation. *Science*. **338**, 1076-1080

17 Zuo, Y. and Steitz, T.A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol Cell*. **58**, 534-540

18 Ebright, R.H. (1993) Transcription activation at Class I CAP-dependent promoters. *Mol Microbiol*. **8**, 797-802

19 Müller-Hill, B. (1996) The *lac* Operon. A Short History of a Genetic Paradigm. Walter de Gruyter, Berlin New York

20 Wilson, C.J., Zhan, H., Swint-Kruse, L. and Matthews, K.S. (2007) The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell Mol Life Sci*. **64**, 3-16

21 Oehler, S., Eismann, E.R., Kramer, H. and Muller-Hill, B. (1990) The three operators of the *lac* operon cooperate in repression. *Embo J*. **9**, 973-979

22 Kramer, H., Niemoller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B. and Muller-Hill, B. (1987) *lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators. *Embo j*. **6**, 1481-1491

23 Oehler, S., Amouyal, M., Kolkhof, P., von Wilcken-Bergmann, B. and Muller-Hill, B. (1994) Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. *Embo J*. **13**, 3348-3355
24 Siebenlist, U., Simpson, R.B. and Gilbert, W. (1980) *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell*. **20**, 269-281
25 Hook-Barnard, I.G. and Hinton, D.M. (2007) Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul Syst Bio*. **1**, 275-293
26 Mitchell, J.E., Zheng, D., Busby, S.J. and Minchin, S.D. (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res*. **31**, 4689-4695
27 Lodge, J., Fear, J., Busby, S., Gunasekaran, P. and Kamini, N.R. (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol Lett*. **74**, 271-276
28 Sadler, J.R., Sasmor, H. and Betz, J.L. (1983) A perfectly symmetric *lac* operator binds the *lac* repressor very tightly. *Proc Natl Acad Sci U S A*. **80**, 6785-6789
29 Simons, A., Tils, D., von Wilcken-Bergmann, B. and Muller-Hill, B. (1984) Possible ideal *lac* operator: *Escherichia coli lac* operator-like sequences from eukaryotic genomes lack the central G X C pair. *Proc Natl Acad Sci U S A*. **81**, 1624-1628
30 Albiniak, A.M., Matos, C.F., Branston, S.D., Freedman, R.B., Keshavarz-Moore, E. and Robinson, C. (2013) High-level secretion of a recombinant protein to the culture medium with a *Bacillus subtilis* twin-arginine translocation system in *Escherichia coli*. *Febs j*. **280**, 3810-3821
31 Alanen, H.I., Walker, K.L., Lourdes Velez Suberbie, M., Matos, C.F., Bonisch, S., Freedman, R.B., Keshavarz-Moore, E., Ruddock, L.W. and Robinson, C. (2015) Efficient export of human growth hormone, interferon alpha2b and antibody fragments to the periplasm by the *Escherichia coli* Tat pathway in the absence of prior disulfide bond formation. *Biochim Biophys Acta*. **1853**, 756-763
32 Bryksin, A.V. and Matsumura, I. (2010) Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques*. **48**, 463-465
33 Sambrook, J. and Russell, D.W. (2001) Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
34 Sarkar, G. and Sommer, S.S. (1990) The "megaprimer" method of site-directed mutagenesis. *Biotechniques*. **8**, 404-407
35 Robinson, C., Matos, C.F., Beck, D., Ren, C., Lawrence, J., Vasisht, N. and Mendel, S. (2011) Transport and proofreading of proteins by the twin-arginine translocation (Tat) system in bacteria. *Biochim Biophys Acta*. **1808**, 876-884
36 Makoff, A.J. and Oxer, M.D. (1991) High level heterologous expression in *E. coli* using mutant forms of the *lac* promoter. *Nucleic Acids Res*. **19**, 2417-2421
37 Balasundaram, B., Harrison, S. and Bracewell, D.G. (2009) Advances in product release strategies and impact on bioprocess design. *Trends Biotechnol*. **27**, 477-485
38 Weiss, D.S., Chen, J.C., Ghigo, J.M., Boyd, D. and Beckwith, J. (1999) Localization of FtsI (PBP3) to the septal ring requires its membrane anchor, the Z ring, FtsA, FtsQ, and FtsL. *J Bacteriol*. **181**, 508-520
39 Garza de Leon, F., Sellars, L., Stracy, M., Busby, S.J.W. and Kapanidis, A.N. (2017) Tracking Low-Copy Transcription Factors in Living Bacteria: The Case of the *lac* Repressor. *Biophys J*. **112**, 1316-1327
40 Browning, D.F. and Busby, S.J. (2016) Local and global regulation of transcription initiation in bacteria. *Nat Rev Microbiol*. **14**, 638-650
41 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*. **33**, 103-119
42 Browning, D.F. and Busby, S.J. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol*. **2**, 57-65
43 Browning, D.F., Cole, J.A. and Busby, S.J. (2008) Regulation by nucleoid-associated proteins at the *Escherichia coli nir* operon promoter. *J Bacteriol*. **190**, 7258-7267

44 Dykxhoorn, D.M., St Pierre, R. and Linn, T. (1996) A set of compatible tac promoter expression vectors. *Gene*. **177**, 133-136

45 Browning, D.F., Matthews, S.A., Rossiter, A.E., Sevastsyanovich, Y.R., Jeeves, M., Mason, J.L., Wells, T.J., Wardius, C.A., Knowles, T.J., Cunningham, A.F., Bavro, V.N., Overduin, M. and Henderson, I.R. (2013) Mutational and topological analysis of the *Escherichia coli* BamA protein. *PLoS One*. **8**, e84512

## FIGURE LEGENDS

**Figure 1. Transcript initiation at bacterial promoters and regulation at the *lac* promoter**
**(A)** Bacterial promoter elements. A sketch showing the organization of different promoter elements at a model bacterial promoter with respect to the transcript start site (+1). Elements are denoted by coloured rectangles with their consensus sequence motifs indicated. Note that bacterial promoters usually contain a selection of these elements but rarely contain all of them [40].
**(B)** Sketch illustrating the key interactions between different domains of the RNAP sigma and alpha subunits and different promoter elements in the initial "closed" complex at a promoter. Note that the DNA duplex remains double-stranded and that the contacts result in Sigma Domain 2 being positioned adjacent to the promoter -10 element.
**(C)** Sketch illustrating the key interactions between different RNAP determinants and different promoter elements in the transcriptionally-competent "open" complex at a promoter. Note that the DNA duplex around position +1 is unwound, with Sigma Domain 2 making specific base contacts with the single-stranded non-template strand of the promoter -10 element, thereby permitting the single-stranded template strand to enter the RNAP active site.
**(D)** A schematic representation of the *E. coli lac* operon promoter. The -35 and -10 promoter elements are shown as green and yellow boxes, respectively, the *lac* operator sequences (O1 and O3) are red boxes and the CRP binding site is shown by inverted arrows. The position of each element is given with respect to the start site of transcription (+1). The LacI tetramer binds to the O1 and O3 operators to form a repression loop, silencing lac operon expression (-ve), whilst CRP activates transcription (+ve).

**Figure 2. Organisation of *lac*, *tac* and *lac*UV5 promoter constructs**
**(A)** The panel shows schematic representations of the *lac* O3O1*, lac*UV5 O3O1 and *tac* O1 promoters and the important elements involved in their regulation. All numbering is in relation to the promoter transcription start (+1), which is indicated by a bent arrow. The O1 and O3 operator sequences, which bind the LacI repressor, are indicated by red boxes, the -35 promoter elements by green boxes, and the CRP site, within the *lac*UV5 and *lac* promoters, is represented by inverted arrows. The *plac* -10 promoter element is shown as a yellow box, whilst the *plac*UV5 and *ptac* consensus -10 elements are gold. The sequence of the O3, O1 and Oid operators, and the site at which different operator sequences were introduced into the *tac* O1, *lac*UV5 O3O1 and *lac* O3O1 promoters is shown by an arrow.
**(B)** The panel shows the base sequence of DNA fragments carrying the *lac* O3O1*, lac*UV5 O3O1 and *tac* O1 promoters. All promoter fragments carry the relevant sequences from -92 to +38, in relation to the transcription start, and possess terminal EcoRI and HindIII sites for cloning into the *lacZ* expression vector, pRW50 [27]. The O1 and O3 operator sequences are highlighted in grey, CRP binding targets are underlined and the -35 and -10 promoter elements are in bold and underlined, having also been aligned to the relevant consensus sequence for each element [40]. Note that the *lac*UV5 promoter only differs from the wild-type *lac* promoter by carrying a consensus -10 element (*i.e.* the p8A and p9A substitutions) [24].

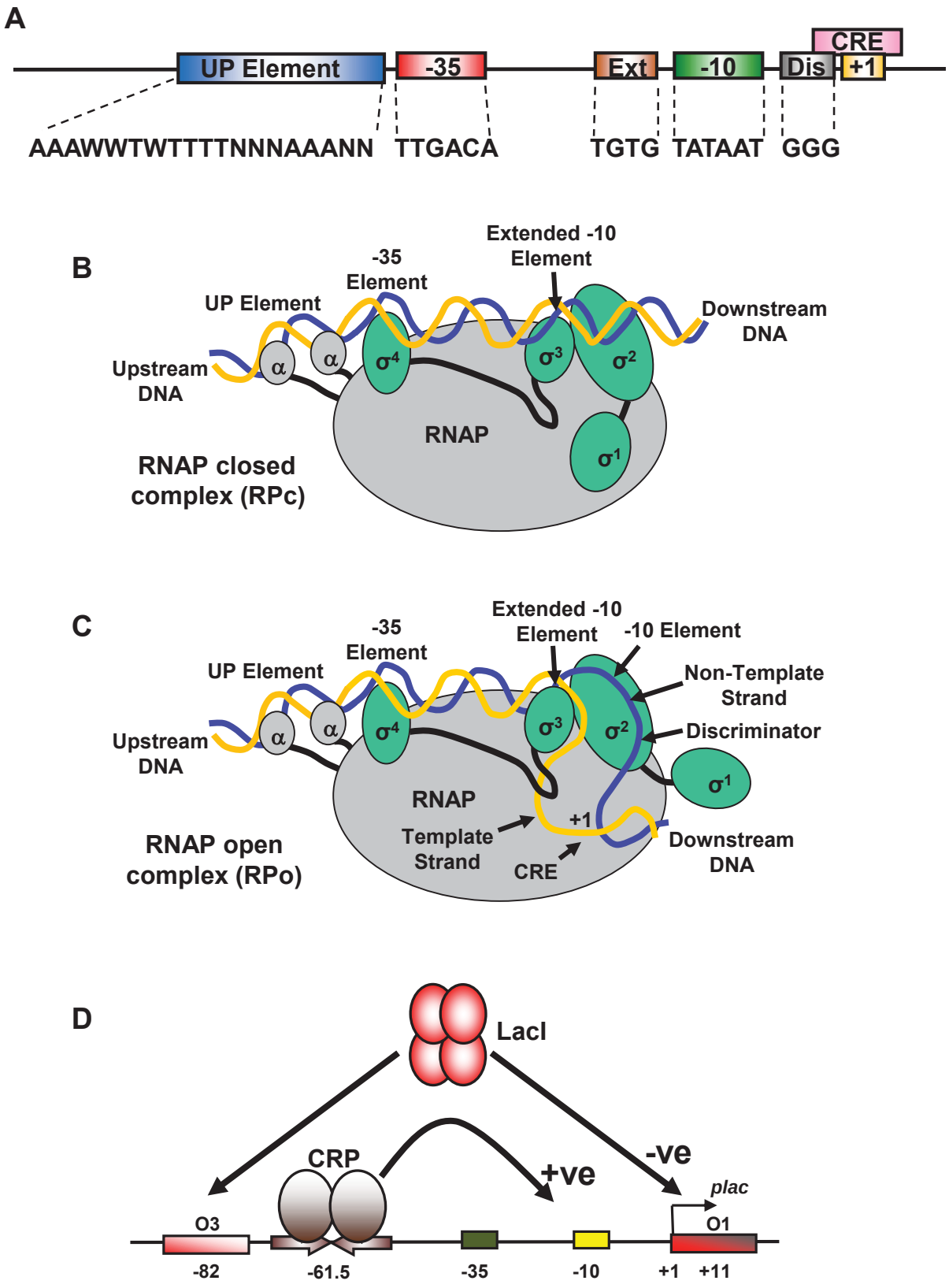
**Figure 3. Expression levels of engineered *ptac* and *plac* promoter derivatives**
The figure illustrates measured β-galactosidase activities of *E. coli* K-12 strain JM109 (Δ*lac*, *lacI*$^{q}$) cells carrying pRW50 [27, 41] containing different promoter fragments. Promoter fragments were generated using PCR and cloned into pRW50 to create *lacZ* transcriptional fusions, using standard techniques [27, 33]. Panels **(A)**, **(B)** and **(C)** show the effect of introducing the O3, O1 and Oid operator sequences into the upstream region of the *tac* O1, *lac*UV5 O3O1 and *lac* O3O1 promoter fragments, respectively, as displayed in Figure 2. **(D)** The panel details the effect of introducing point mutations into the -10 and -35 elements of various *lac* promoter derivatives (*i.e. lac* O3O1, *lac* O1O1, *lac*UV5 O3O1, *lac*UV5 O1O1) to

improve these regions in relation to the extended -10 and -35 consensus sequences (see Figure 2) [42]. By convention, locations are labelled in relation to the *plac* transcript start point (+1). Note that the *lac* O1O1 D19 promoter carries a base pair deletion at position -19. The fold increase in expression, in comparison to the weakest promoter (*lac* O1O1) is indicated for each promoter in the presence of IPTG. In all panels, JM109 cells were grown in LB medium until mid-exponential phase, in the presence or absence of 1 mM IPTG. β-galactosidase activities were determined using o-nitrophenyl-β-D-galactopyranose (ONPG) and a Miller protocol, as in our previous work [43], and are expressed as nmol of ONPG hydrolysed $min^{-1}$ $mg^{-1}$ dry cell mass. Each activity is the average of three independent determinations and the error bars represent the standard deviation of values.

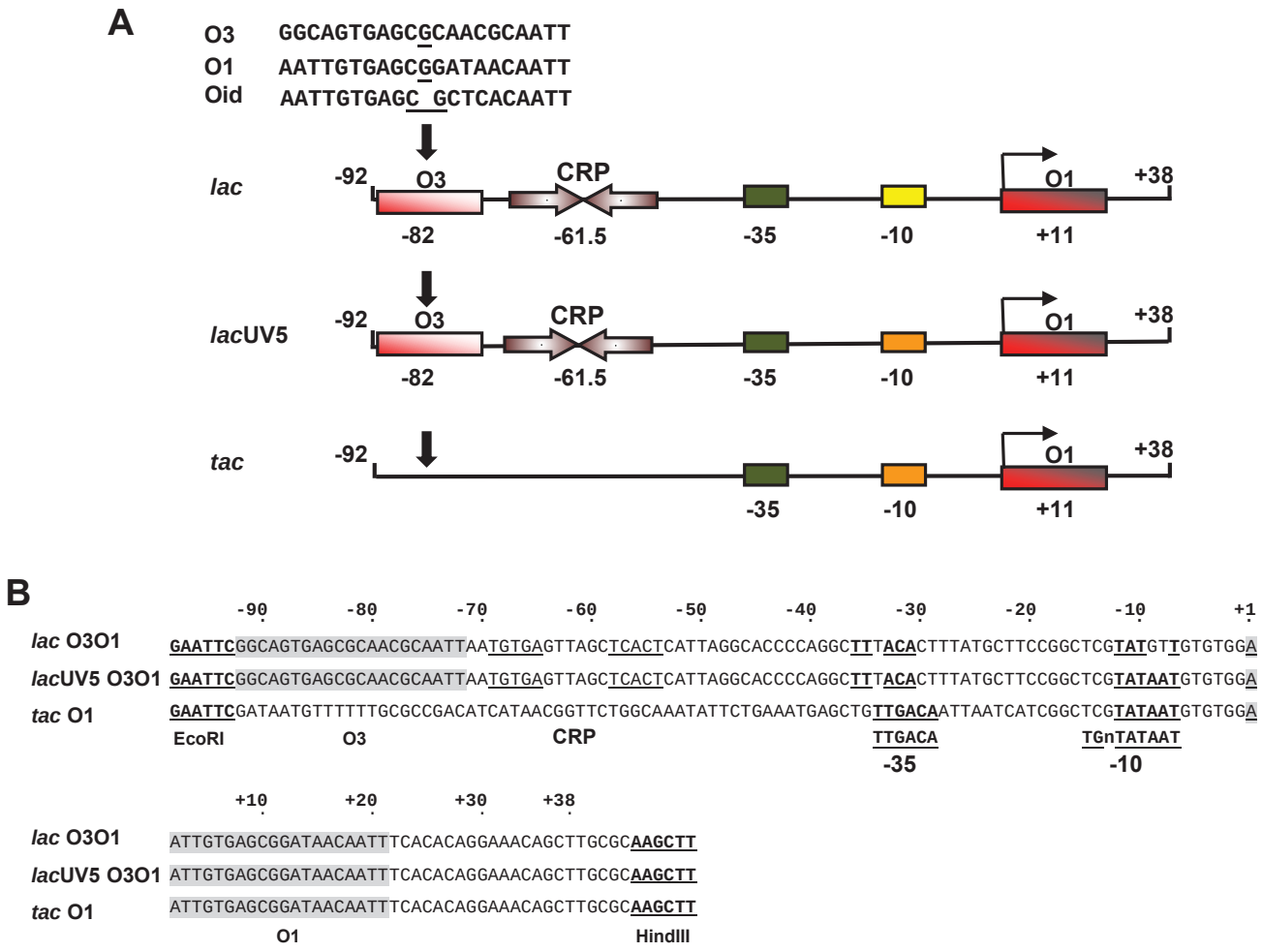
**Figure 4. The expression of TorA-protein fusions can be set to different levels with PAR promoter constructs**

**(A)** Expression from the PAR promoter constructs. The panel shows the β-galactosidase activities measured in the *E. coli* K-12 strain JM109 (Δ*lac*, *lacI*$^q$) containing the *lacZ* expression vector pRW50 carrying various PAR promoters. PAR promoters (PAR1 to PAR8) were selected from the various *tac* and *lac* promoters detailed in Figure 3. β-galactosidase activities are expressed as nmol of ONPG hydrolysed $min^{-1}$ $mg^{-1}$ dry cell mass. Each activity is the average of three independent determinations and standard deviations are shown for all data points. **(B)** Expression of a TorA-GFP-6His protein fusion in BL21 cells (Novagen), using pGFP-TorA based vectors [30]. PAR promoters, PAR4 to PAR6 and PAR8 were introduced into the low-copy-number vector pGFP-TorA, such that each promoter drives the expression of a *torA-gfp-6his* fusion. Note that pGFP-TorA is a derivative of expression vector pEXT22 [44], which has been included as an empty vector control. **(C)** Expression of a TorA-hGH-6His protein fusion in BL21 cells, using pHAK1 based vectors [31]. PAR promoters, PAR5 to PAR8 were introduced into the high-copy-number vector pHAK1 such that each promoter drives the expression of a *torA-hGH-6his* fusion. Note that the *torA* signal sequence directs the GFP-6His and hGH-6His moieties to the Tat translocon for periplasmic targeting [7, 30, 31]. Cells were grown in LB medium until an $OD_{600}$ of ~0.4 when RPP was induced by the addition of 1 mM IPTG for 3 hrs. In **(B)** and **(C)** expression of TorA-GFP-6His and TorA-hGH-6His was analysed by Western blotting (upper panel), using either anti-GFP antibody (Sigma) or polyclonal anti-hGH antibody, respectively, and a Coomassie blue strained SDS-PAGE gel (lower panel), using normalised total cell protein from BL21 cells, as detailed in [45]. Samples were calibrated by loading Page Ruler Plus prestained markers (Thermo Scientific). Note that, in the Western blot in panel **(B),** all lanes were from the same blot and have only been separated to aid presentation of the data.

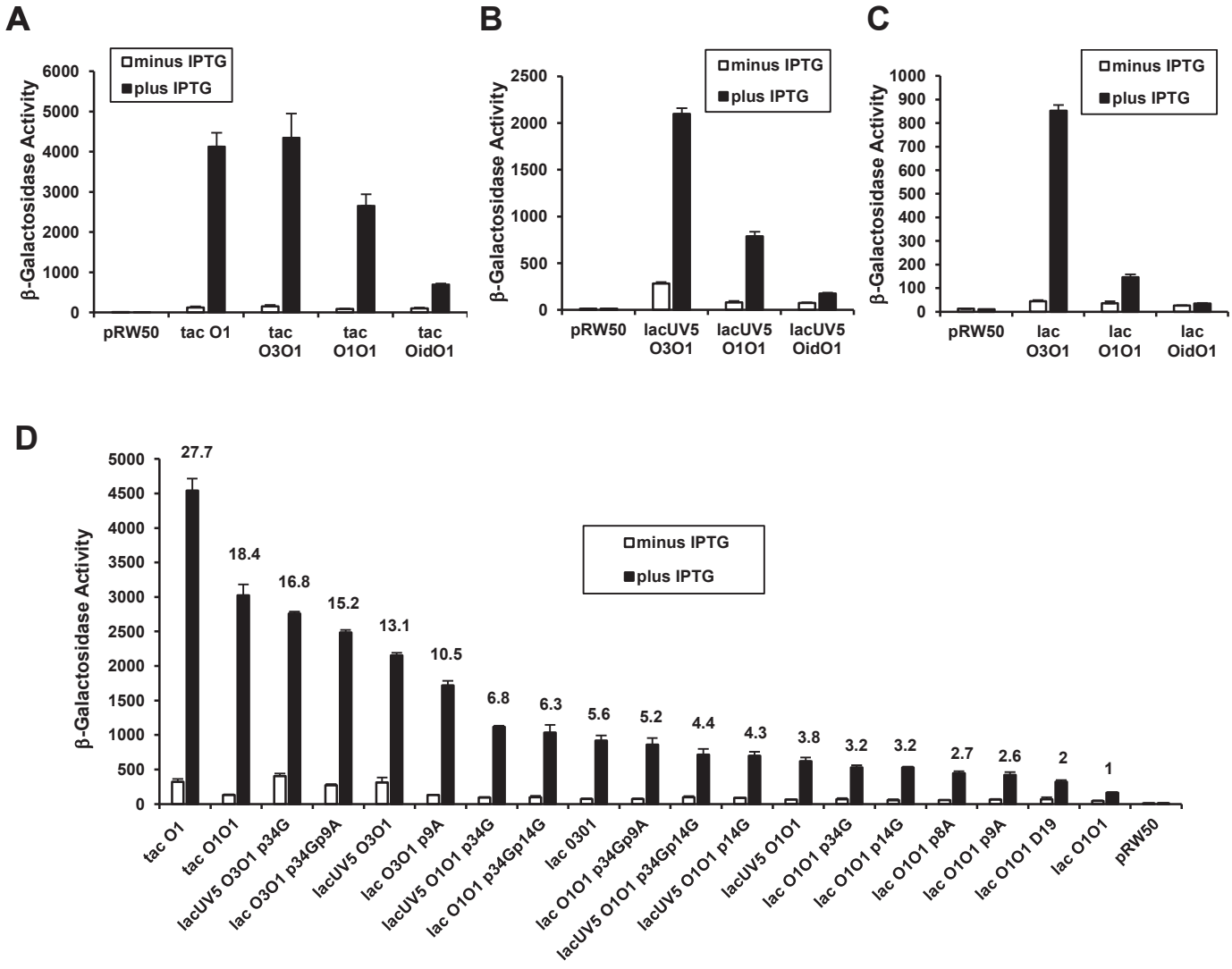
Figure 1


A
UP Element
-35
Ext
-10
Dis
CRE
+1
AAAWWTWTTTTNNNAAANN
TTGACA
TGTG
TATAAT
GGG
B
Extended -10 Element
-35 Element
UP Element
Downstream DNA
Upstream DNA
α
α
σ4
σ3
σ2
RNAP
σ1
RNAP closed complex (RPc)
C
Extended -10 Element
-10 Element
-35 Element
UP Element
Non-Template Strand
Upstream DNA
α
α
σ4
σ3
σ2
Discriminator
σ1
RNAP
+1
Template Strand
CRE
Downstream DNA
RNAP open complex (RPo)
D
LacI
CRP
+ve
-ve
plac
O3
O1
-82
-61.5
-35
-10
+1
+11

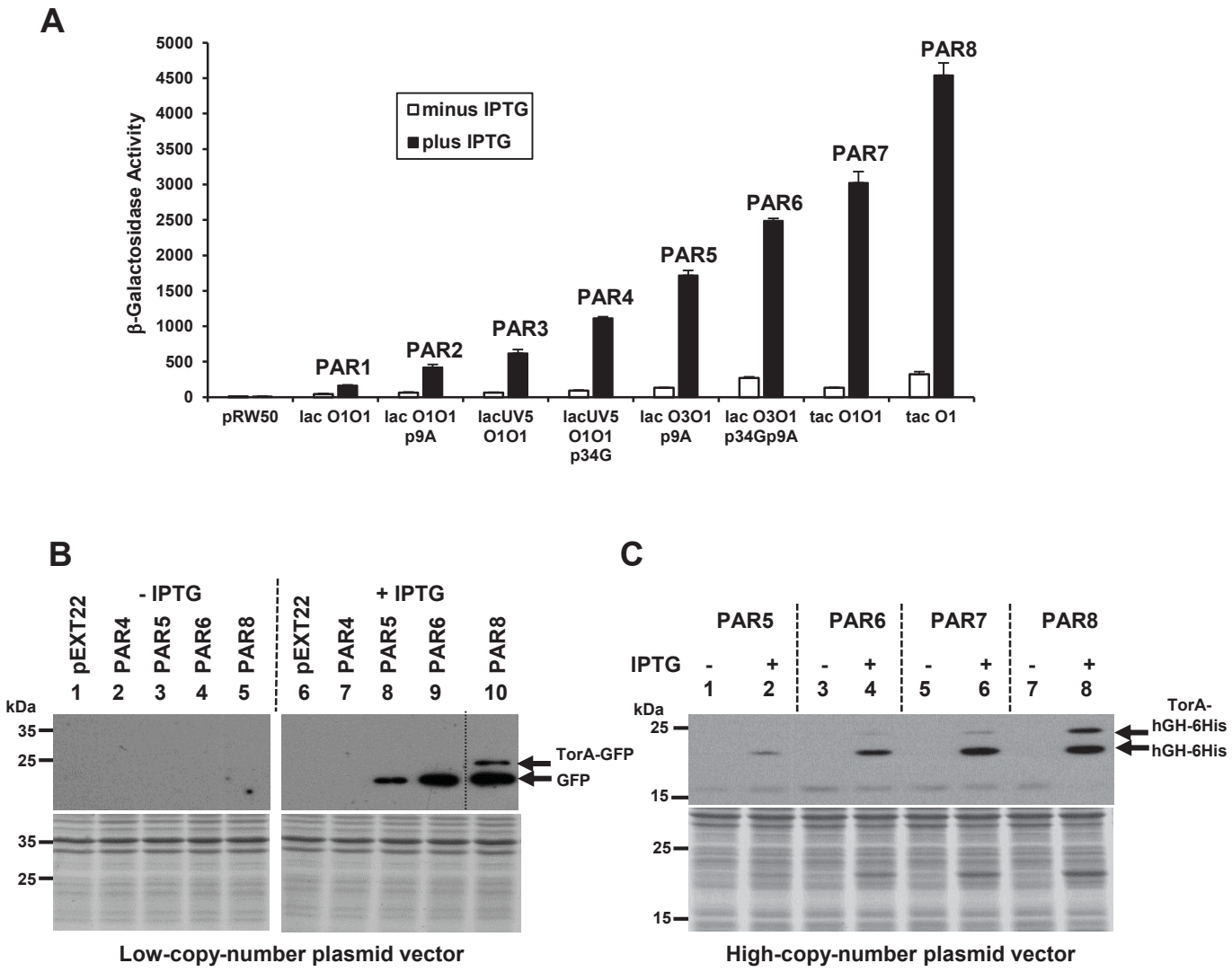
# Figure 2


A
O3 GGCAGTGAGCGCAACGCAATT
O1 AATTGTGAGCGGATAACAATT
Oid AATTGTGAGC GCTCACAATT
lac
lacUV5
tac
-92
O3
CRP
O1
+38
-82
-61.5
-35
-10
+11
B
-90 -80 -70 -60 -50 -40 -30 -20 -10 +1
lac O3O1 GAATTCGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA
lacUV5 O3O1 GAATTCGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGA
tac O1 GAATTCGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGA
EcoRI
O3
CRP
TTGACA
-35
TGnTATAAT
-10
+10 +20 +30 +38
lac O3O1 ATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTTGCGCAAGCTT
lacUV5 O3O1 ATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTTGCGCAAGCTT
tac O1 ATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTTGCGCAAGCTT
O1
HindIII

# Figure 3


A
B
C
D
β-Galactosidase Activity
minus IPTG
plus IPTG
pRW50
tac O1
tac O3O1
tac O1O1
tac OidO1
lacUV5 O3O1
lacUV5 O1O1
lacUV5 OidO1
lac O3O1
lac O1O1
lac OidO1
27.7
18.4
16.8
15.2
13.1
10.5
6.8
6.3
5.6
5.2
4.4
4.3
3.8
3.2
3.2
2.7
2.6
2
1
tac O1
tac O1O1
lacUV5 O3O1 p34G
lac O3O1 p34Gp9A
lacUV5 O3O1
lac O3O1 p9A
lacUV5 O1O1 p34G
lac O1O1 p34Gp14G
lac 0301
lac O1O1 p34Gp9A
lacUV5 O1O1 p34Gp14G
lacUV5 O1O1 p14G
lacUV5 O1O1
lac O1O1 p34G
lac O1O1 p14G
lac O1O1 p8A
lac O1O1 p9A
lac O1O1 D19
lac O1O1
pRW50

# Figure 4


A
β-Galactosidase Activity
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
minus IPTG
plus IPTG
PAR1
PAR2
PAR3
PAR4
PAR5
PAR6
PAR7
PAR8
pRW50
lac O1O1
lac O1O1 p9A
lacUV5 O1O1
lacUV5 O1O1 p34G
lac O3O1 p9A
lac O3O1 p34Gp9A
tac O1O1
tac O1
B
- IPTG
+ IPTG
pEXT22
PAR4
PAR5
PAR6
PAR8
1 2 3 4 5 6 7 8 9 10
kDa
35
25
TorA-GFP
GFP
Low-copy-number plasmid vector
C
PAR5
PAR6
PAR7
PAR8
IPTG - + - + - + - +
1 2 3 4 5 6 7 8
kDa
25
15
TorA-hGH-6His
hGH-6His
High-copy-number plasmid vector