High level heterologous expression in *E.coli* using mutant forms of the *lac* promoter

A.J.Makoff* and M.D.Oxer

Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK

Received January 16, 1991; Revised and Accepted April 12, 1991

ABSTRACT

Single base deletions in the *lac* promoter which reduced the 18bp spacing between the -35 and -10 homology regions to 17bp, increased the strength of the promoter. A single base substitution (T→G) in the -35 region to generate the consensus sequence TTGACA increased the strength further and no longer required a 17bp spacing. The mutated *lac* promoter was as powerful as a shorter form of the *tac* promoter which lacked two AT-rich regions upstream of the -35 region, and expressed the P69 surface antigen (pertactin) of *Bordetella pertussis* to 30–40% total cell protein and tetanus toxin fragment C to 16–20% total cell protein.

INTRODUCTION

Promoters which are recognised by *E. coli* RNA polymerase have two regions of sequence homology which are located approximately 35 and 10 nucleotides upstream of the start of transcription (1,2). Spacing between the two homology regions has been shown to be important for promoter function, the optimum spacing usually being 17 nucleotides (3–6). However, there are clearly other less well-defined features which also affect transcription levels (7,8); in particular, two AT-rich regions upstream of the -35 homology region of the *trp* promoter (9).

A frequently used promoter for heterologous expression is the *tac* promoter, a hybrid between the *trp* and *lac* promoters which combines the utility of induction of the *lac* promoter with the greater strength of the *trp* promoter (10,11). We have successfully used the *tac* promoter to express high levels of a number of heterologous proteins (12–14). In some of these examples, a shorter and weaker form of the *tac* promoter, which lacks the *trp* promoter upstream AT-rich regions, would substitute for the long form with no reduction in expression levels. However, the *lac* promoter was a poor substitute, resulting in much lower levels of expression.

In order to improve the *lac* promoter and to investigate possible causes of its weakness relative to the *tac* promoter, we have made small sequence changes which alter either the -35 region and/or the spacing between the two homology regions. We have examined the effects of these changes on the expression of two heterologous proteins: P69 surface antigen (pertactin) from

Bordetella pertussis and tetanus toxin fragment C. We have previously expressed both proteins in *E. coli*, using the *tac* promoter, and showed that very high levels of pertactin could be produced using the truncated *tac* promoter (14). Expression of fragment C was initially limited by the poor codon bias of the *Clostridium tetani* toxin gene, but when a chemically synthesized gene was used to overcome this problem, mRNA levels became limiting, requiring the full length *tac* promoter for higher expression (12).

MATERIALS AND METHODS

Strain

E. coli strain MM294 (15) was used throughout this work.

Plasmids

Plasmid pPERTac36 has been described elsewhere (14). The *tac* promoter sequence of pPERTac36 was replaced by cloning an oligonucleotide containing the *lacUV5* sequence (16). Plasmid pPERlac1, obtained after screening many transformants, had a single C→T base change 6 nucleotides upstream of the -35 sequence of the *lac* promoter, which created a unique *Bsu36I* site (Fig. 1).

Three oligonucleotides (oligo-1, 2 and 3, Fig. 1) were cloned between the *Bsu36I* and *ApaI* sites of pPERlac1. Plasmid DNA was isolated from a number of transformants and the *lac* promoter region was sequenced. The sequence changes in four plasmids, pPERlac20, 18, 19 and 16 are given in Fig. 1. The same three oligonucleotides were also ligated to the *Bsu36I*-*BamHI* fragment of pPERlac1 and the *ApaI*-*BamHI* fragment of pTETtac11 (12), encoding the fragment C gene, which after transformation gave an equivalent range of plasmids expressing fragment C. Plasmid DNA was also isolated from these transformants and the promoter region sequenced, which identified two plasmids with the same promoter mutations as pPERlac19 and 16.

Oligonucleotides which encode the *trpA* terminator preceded by *EcoRI* and *BamHI* sites, were cloned into pUC19 between the two *HaeII* sites which flank the *lacI* and *lacZ* fragments and the polylinker (17). A plasmid with the *trpA* terminator in the same orientation as in our pAT153-based plasmids was selected and the *EcoRI*-*ApaI* fragment from pPERlac1, encoding the *lacI*

* To whom correspondence should be addressed

gene and *lac* promoter, and the *Apal-BamHI* fragment from pTETlac115 (12), encoding the fragment C gene were cloned into it, to generate pTETlac5 (Fig. 2). Plasmids pTETlac2, 4 and 3 were generated from pTETlac5 and the fragment C plasmid equivalent to pPERlac16 (see above). This was achieved by isolating the small *Bsu36I-SalI* fragment (see Fig. 2) which flanks the *lac* promoter region of the above two plasmids and digesting with *HapII*, which cleaves once within each fragment. The mixed cleaved fragments were then cloned between the *Bsu36I* and *SalI* sites of pTETlac5, and the resulting plasmids were identified by sequencing.

Plasmid pTETtac215 was created from pTETlac5 and pTETtac115, using the *AatII* and *BamHI* sites common to both plasmids. Plasmid pTETtac315 was constructed from pTETlac5 and pTETtac11 (12) and plasmids pTETlac9 and 8 from pTETlac5 and either pPERlac20 or 18 respectively (Fig. 1), in each case using the *PfI*MI and *BglII* sites common to all these plasmids.

All oligonucleotides were synthesized on a Pharmacia Gene Assembler and the resulting plasmids confirmed by sequencing as previously described (12).

Induction of cultures

Cultures were induced, analysed and expression levels quantified as described previously (12,18,19).

Northern blots

RNA was prepared, separated on an agarose gel, blotted, hybridized, autoradiographed and quantified as described previously (12). Similarly treated RNA size markers (BRL) were included in the gel and after blotting, the appropriate lane on the membrane was excised and stained with methylene blue. A 927bp *ClaI* fragment from pTETtac11 was used as a probe and labelled as previously described (12).

RESULTS

Expression of pertactin

The *tac* promoter in the plasmid pPERTac36 (14) was replaced by the *lac* promoter to generate the plasmid pPERlac1, which on induction yielded much lower yields of pertactin (Fig. 3). A mixed oligonucleotide (Fig. 1) was cloned between the *Bsu36I* and *Apal* sites of pPERlac1 to produce a series of plasmids. The two adjacent degenerate positions of oligo-1 replace three nucleotides in pPERlac1, shown in Fig. 1 between the two dotted lines. This would be expected to yield a single base deletion, which in 75% of the plasmids would be at one of three positions between the -35 and -10 regions of the *lac* promoter. Additionally, because of the third degenerate position, 50% of the plasmids should bear a T→G base change in the -35 region, resulting in the consensus TTGACA.

A number of these plasmids were selected and the *lac* promoter regions were sequenced. Four plasmids: pPERlac20, 18, 19 and 16 (Fig. 1) were chosen for further study. Inductions of cells containing pPERlac1, pPERTac36 and the new plasmids were compared by SDS-PAGE of cell extracts. As shown in Fig. 3, all four plasmids produced more pertactin than pPERlac1, and at levels comparable to pPERTac36, which was confirmed by densitometer scanning of gels of many induced cell extracts (30–40% total cell protein, data not shown).

Fig. 1. Sequence changes in the *lac* promoter of various plasmids. The two homology regions are shown in bold italics. Restriction enzyme sites are underlined. The two vertical dotted lines indicate the two nucleotides which flank the three base region, reduced to two bases by cloning the mixed oligonucleotide. The Δ symbols indicate a single base deletion.

Fig. 2. Map of pTETlac5. The *HaeII* fragment containing the β -lactamase gene (*Amp^r*) and the replication origin (*ori*) is from pUC19. The *Bsu36I-Apal* fragment containing the *lac* promoter (*P_{lac}*) is shown in more detail in Fig. 1. The tetanus toxin fragment C gene (*frag C*) is followed by the *trpA* terminator (*term*).

Expression of tetanus toxin fragment C

Although the four plasmids with a mutated *lac* promoter are all as effective as pPERTac36 in giving high levels of pertactin, it does not necessarily follow that the four mutant *lac* promoters are equivalent in strength to the *tac* promoter. In order to discriminate between the different promoter strengths of the mutant *lac* promoters and two forms of the *tac* promoter, use of a gene whose expression is limited by these higher mRNA levels would be necessary. This requirement is satisfied by the expression of the synthetic codon-optimized gene for tetanus toxin fragment C (12). Preliminary experiments using the *lac* promoter to express fragment C suggested that the levels were insufficient to allow accurate estimation of expression to be determined by SDS-PAGE followed by densitometer scanning. In order to overcome this problem, fragment C expression levels were

Fig. 3. Coomassie blue stained SDS-polyacrylamide gel of extracts of induced cells expressing pertactin. The numbers on each track: 1, 20, 18, 19, 16 refer to the pPERlac plasmids, 36 refers to pPERtac36 and C refers to the control track.

Fig. 4. Coomassie blue stained SDS-polyacrylamide gel of extracts of induced cells expressing tetanus toxin fragment C. The numbers on each track: 5, 4, 3, 2 refer to the pTETlac plasmids, 215 and 315 refer to the pTETtac plasmids and C refers to the control track.

increased by replacing the pAT153 replicon by the higher copy number replicon of pUC19 (17).

Plasmid pTETlac5 (Fig. 2) is such a plasmid, containing the pUC19 replicon and fragment C under the control of the *lac* promoter. Plasmid pTETlac5 was used to construct pTETtac215,

Table 1. Levels of fragment C and of its specific mRNA expressed by various plasmids

Plasmid	Promoter	Expressed Percentage Cell Protein	Fragment C Relative Levels*	mRNA Relative Levels*
pTETlac5	<i>lac</i>	7.1 ± 0.5 (8)	1.0 ± 0.1	1.0 ± 0.3 (4)
pTETlac9	<i>lac</i> (ΔG)	14.6 ± 2.1 (4)	2.1 ± 0.3	ND
pTETlac8	<i>lac</i> (ΔC)	14.1 ± 1.6 (4)	2.0 ± 0.2	ND
pTETlac2	<i>lac</i> (ΔT)	12.9 ± 1.1 (8)	1.8 ± 0.2	1.4 ± 0.4 (4)
pTETlac4	<i>lac</i> (T→G)	17.3 ± 1.0 (4)	2.4 ± 0.1	2.5 ± 0.4 (4)
pTETlac3	<i>lac</i> (T→G,ΔT)	18.3 ± 1.3 (4)	2.6 ± 0.2	3.2 ± 0.9 (4)
pTETtac315	short <i>tac</i>	20.7 ± 2.5 (4)	2.9 ± 0.4	2.6 ± 0.4 (4)
pTETtac215	long <i>tac</i>	24.2 ± 1.7 (4)	3.4 ± 0.2	4.8 ± 0.7 (4)

* Levels are relative to those of pTETlac5.

ND = Not determined.

The number of determinations for each plasmid is shown in brackets.

Each figure is the mean expression level of fragment C ± the standard error of the mean.

Fig. 5. Northern blot of RNA isolated from induced cultures. The number on each track refers to the plasmid involved as in Fig. 4. The sizes are given by RNA markers and are in kilo nucleotides.

containing the full-length *tac* promoter, pTETtac315 where the shorter version of the *tac* promoter was used, and plasmids pTETlac9, 8, 2 and 3, using the same *lac* promoter mutations as pPERlac20, 18, 19 and 16. Plasmid pTETlac4 was also derived from pTETlac5 (Fig. 1), enabling an assessment to be made of the effect of the consensus TTGACA without a single base deletion in the spacer region.

Expression of fragment C was compared in induced cultures containing some of the above plasmids (Fig. 4). The expression levels were quantitated by densitometer scanning of several independent inductions for each plasmid including pTETlac9 and 8 (Table 1). This confirms the result with the pertactin plasmids showing that reducing the size of the spacer region of the *lac* promoter to 17 nucleotides considerably improved promoter strength independently of the base deleted. It also shows that replacing the -35 sequence by the consensus TTGACA increases

the promoter strength still further, to that of the shorter *tac* promoter, irrespective of whether the 18 or 17 nucleotide spacing is used.

Estimates of mRNA levels

In order to assess promoter strength more directly, Northern blots were carried out on RNA isolated from induced cells (Fig. 5). The levels of mRNA was quantitated by densitometer scanning of Northern blot autoradiographs of four independent inductions for each plasmid (Table 1). The mRNA levels show that the full-length *tac* promoter is approximately five fold stronger than the *lac* promoter, consistent with previous comparisons (7,10,11). The mRNA levels clearly follow the same trend as the expressed protein levels and confirm those results.

DISCUSSION

We have used the production of two heterologous proteins *in vivo* as a measure of promoter strength. Estimates using expression of the pertactin gene were clearly unsatisfactory since its expression levels were not limited by mRNA levels in the range achieved by the vectors used in this work. By contrast, expression of the synthetic gene for tetanus toxin fragment C was able to discriminate between the mutant *lac* and *tac* promoters. Northern blots confirmed that the different levels of fragment C reflect the different mRNA levels. However, for the relative steady state mRNA levels to reflect promoter strengths, we must assume that (a) the rate of degradation of the specific transcripts is the same, (b) the copy number of the plasmids is the same and (c) the plasmids are equally stable. (a) The Northern blot reveals a number of smaller mRNA bands in addition to the full-length transcript, which are probably due to degradation both *in vivo* and during the RNA extraction. They show approximate proportionality with the major band which makes it unlikely that there are gross discrepancies between the rates of degradation of each transcript. However, some degradation products might be too short-lived to be detected by this means. (b) It is possible that either high rates of transcription or high levels of heterologous proteins might lead to a decrease in copy number. We cannot rule out this possibility. (c) It is known that high levels of heterologous proteins can sometimes result in plasmid loss. We have found no evidence for plasmid loss during any of the inductions described in this paper (data not shown).

When the *tac* promoter was initially constructed, it was not clear what features were responsible for making it more powerful than the *lac* promoter (10). One possibility was that it resulted solely from the combination of both the consensus -35 sequence (TTGACA) and the consensus -10 sequence (TATAAT). However, Russell and Bennett showed that a hybrid *tet/lac* promoter which also brought together the two consensus sequences, separated by 16 nucleotides as in the *tac* promoter, is no more powerful than the *lac* promoter (20). Another possibility is that it is due to the presence of upstream AT-rich sequences in both the *trp* and *tac* promoters, removal of which reduces the strength of the *trp* promoter (9). These AT-rich sequences clearly do contribute to the strength of the *tac* promoter (12), but even without them the *tac* promoter is still stronger than the *lac* promoter. We have shown here that the truncated *tac* promoter is 2–3 times stronger than the *lac* promoter which is in agreement with the comparison made by Russell and Bennett (20), whose *tac* promoter also lacks the upstream sequences.

In a recent study *in vitro*, comparison of the promoter strengths

of all possible single base substitutions in the -35 region of the *lac* promoter showed that the mutant with the consensus -35 region was the strongest (21). Our experiments *in vivo* are consistent with this result, showing a 2–3 fold stimulation in promoter strength when the *lac* promoter is modified in this way. We further showed that this mutant is indistinguishable in strength from the truncated *tac* promoter. The full length *tac* promoter therefore owes its greater strength, when compared with the *lac* promoter, to both the combination of the two consensus sequences and to the upstream AT-rich sequences.

Many studies have examined the effects of alterations in the spacing between the -35 and -10 homology regions. *In vitro* studies have shown that for the *lac* promoter, the optimum spacing is 17bp (3). The results of our experiments *in vivo* are consistent with these findings. This agreement contrasts with discrepancies in similar studies carried out on the *tac* promoter which gave different optima of 17 or 16bp, depending on whether promoter strength was measured respectively *in vitro* or *in vivo* (5,6).

Stefano and Gralla (3) suggested that the 1bp deletion in the 18bp spacer of the *lac* promoter increased promoter strength by facilitating simultaneous contact between the two homology regions and RNA polymerase. Since we have shown that a similar deletion has no effect if the *lac* promoter contains the consensus -35 sequence, this might imply that simultaneous contact is more easily achieved when both consensus regions are present. Another possibility is that creating the consensus -35 sequence in the *lac* promoter alters the optimum spacing (from 17 to 17.5bp) rather than reducing its importance. This could result from a change in the way that the -35 region binds to RNA polymerase and consequently alters the optimum relative orientation of the two homology regions. Although two other promoters containing both consensus sequences show a preference for a spacing of 17bp over 18bp (4–6), comparison with the *lac* promoter is complicated by their different spacing sequences.

We have shown that with a single base change, the *lac* promoter is capable of expressing very high levels of two different heterologous proteins: *B. pertussis* pertactin and tetanus toxin fragment C to 30–40% and 16–20% total cell protein respectively. We would also expect it to express γ -interferon to 30–40% total cell protein, since that level was achieved by our plasmids using the truncated *tac* promoter (13). This modified *lac* promoter is clearly an acceptable alternative to the *tac* promoter for high level expression.

ACKNOWLEDGEMENTS

We would like to thank I.G. Charles for his suggestion of using pUC-based plasmids and J.J. Clare and M.A. Romanos for helpful comments about the manuscript.

REFERENCES

1. Rosenberg, M. and Court, D. (1979) *Ann. Rev. Genet.* **13**, 319–353.
2. Siebenlist, U., Simpson, R. B. and Gilbert, W. (1980) *Cell* **20**, 269–281.
3. Stefano, J. E. and Gralla, J. D. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1069–1072.
4. Aoyama, T., Takanami, M., Ohtsuka, E., Taniyama, Y., Marumoto, R., Sato, H. and Ikehara, M. (1983) *Nucl. Acids Res.* **11**, 5855–5864.
5. Mulligan, M. E., Brosius, J. and McClure, W. R. (1985) *J. Biol. Chem.* **260**, 3529–3538.
6. Brosius, J., Erfle, M. and Storella, J. (1985) *J. Biol. Chem.* **260**, 3539–3541.
7. Deuschle, U., Kammerer, W., Gentz, R. and Bujard, H. (1986) *EMBO J.* **11**, 2987–2994.

8. Kammerer, W., Deuschle, U., Gentz, R. and Bujard, H. (1986) *EMBO J.* **11**, 2995–3000.
9. Nishi, T. and Itoh, S. (1986) *Gene* **44**, 29–36.
10. de Boer, H. A., Comstock, L. J. and Vasser, M. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 21–25.
11. Amann, E., Brosius, J. and Ptashne, M. (1983) *Gene* **25**, 167–178.
12. Makoff, A. J., Oxer, M. D., Romanos, M. A., Fairweather, N. F. and Ballantine, S. (1989) *Nucl. Acids Res.* **17**, 10191–10202.
13. Makoff, A. J. and Smallwood, A. E. (1990) *Nucl. Acids Res.* **18**, 1711–1718.
14. Makoff, A. J., Oxer, M. D., Ballantine, S. P., Fairweather, N. F. and Charles, I. G. (1990) *Bio/Technology* **8**, 1030–1033.
15. Meselson, M. and Yuan, R. (1968) *Nature (London)* **217**, 1110–1114.
16. Fuller, F. (1982) *Gene* **19**, 43–54.
17. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103–119.
18. Makoff, A. J., Parry, N. and Dicken, L. P. (1989) *J. Gen. Microbiol.* **135**, 11–24.
19. Makoff, A. J., Ballantine, S. P., Smallwood, A. E. and Fairweather, N. F. (1989) *Bio/Technology* **7**, 1043–1046.
20. Russell, D. R. and Bennett, G. N. (1982) *Gene* **20**, 231–243.
21. Kobayashi, M., Nagata, K. and Ishihama, A. (1990) *Nucl. Acids Res.* **18**, 7367–7372.