A plasmid vector that allows fusion of the *Escherichia coli* galactokinase gene to the translation startpoint of other genes

Ann Kelsall, Christine Evans and Stephen Busby*

University of Birmingham, Department of Biochemistry, PO Box 363, Birmingham B15 2TT, England

Received 8 October 1984

By genetic manipulation of cloned *Escherichia coli* galactose operon DNA, we have constructed a new plasmid in which the N terminal segment of the *galK* gene is replaced by the N terminal of the *galE* gene. This plasmid encodes a hybrid protein that confers a Gal K^+ phenotype on host cells: differences in initiation at the *galE* translation start point cause different phenotypes. The plasmid has unique restriction sites at the junction of the *galE* and *galK* gene segments and thus can be used to replace the N terminal of *galK* with any other translation start.

Escherichia coli *Plasmid vector* *Galactokinase gene* *Protein fusion* Gal*E*-Gal*K hybrid* *Translation initiation*

1. INTRODUCTION

Gene fusion techniques are widely used to study DNA sequences that promote transcription in procaryotes [1]. Any selectable gene can be inserted downstream of promoter sequences, so that the expression of this gene (the 'indicator' gene) can be used to monitor promoter activity [2–4]. Although a wide variety of vectors can be used, the most common carry either the β-galactosidase (*lacZ*) or galactokinase (*galK*) genes as indicators [1,5].

Gene fusions can also be used to study translation initiation signals: to do this, the indicator gene must be fused to the translation start point under examination. Although this may affect the activity of the indicator gene, it is well known, in the case of β-galactosidase, that the N-terminal can be replaced without destroying enzyme activity [6]: plasmid vectors where the *lacZ* gene can be fused to any foreign N-terminal segment have been developed to study translation initiation signals [7]. Here we report that the N-terminal of galactokinase can be replaced and that the resultant hybrid protein has galactokinase activity. We describe a plasmid vector in which galactokinase is fused to a segment coding for the N-terminal of the *galE* gene and we show that this can be used to create phenotypes that reflect the frequency of translation initiation.

2. EXPERIMENTAL

The starting plasmid, pAA113, has been reported previously [8]. It carries the entire *E. coli gal* operon with a unique *Bst*EII site just after the start of the *galE* gene and a single *Pvu*II site located 170 bp upstream of the *galK* gene [9] (see fig.1a). This plasmid was linearised with *Pvu*II and deletions of different lengths, some cutting into the *galK* gene, were made using *Bal*31 exonuclease (fig.1b). Exonuclease-treated DNA was then restricted with *Bst*EII and the *Bst*EII-end carrying the N-terminal of *galE* was ligated to the end carrying *galK* that had been digested with *Bal*31

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

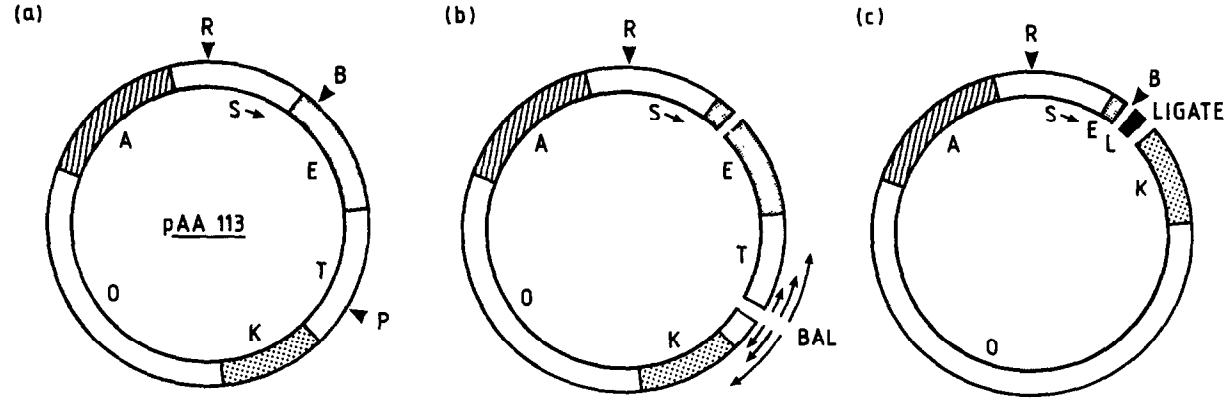

Fig.1. (a) Diagram of pAA113. The plasmid carries the pBR322 replication origin (o) and the galactose operon transcription start (S). Different genes are variously shaded and are labelled: E, UDP galactose 4′-epimerase; T, galactose-1 phosphate transferase; K, galactokinase; and A, penicillinase. The sites of the unique *Eco*RI (R), *Bst*EII (B) and *Pvu*II (P) sites are marked. The plasmid, which is not drawn to scale, consists of 6000 bp. The distance *Eco*RI–*Bst*EII is 130 bp, and *Bst*EII–*Pvu*II is 1900 bp. The *gal*E, T and K genes are each approx. 1100 bp in length [8]. (b) pAA113 was cleaved at the *Pvu*II site, treated with *Bal*31 exonuclease to make deletions of various lengths and subsequently cut with *Bst*EII. (c) The longer DNA fragment was ligated in the presence of a 10 bp linker fragment (L). In the resulting plasmid the N-terminal segment of the *gal*E gene (E) is joined via the linker to a point before or in the *gal*K gene depending on the length of sequence deleted by *Bal*31. In the case of pAA121, the *gal*E and K genes were fused in phase. The *gal* promoter and the segment of *gal*E can be removed using the unique *Eco*RI(R) and *Bst*EII(B) sites.

(fig.1b,c). The result was to fuse the start of *gal*E to different points either before or in the *gal*K gene (fig.1c). To retain restriction sites at the junction of the *gal*E and K genes, the ligation was performed in the presence of a 10 bp *Bst*EII–*Hin*dIII linker fragment isolated from plasmid pAA187 [8].

After this construction, performed using standard recombinant DNA techniques [10], MI101 *Δgal* cells [8] were transformed with the ligated DNA and the plasmid in purified single colonies was prepared and restricted [10]. The restriction pattern confirmed that the construction had proceeded correctly and showed the approximate length of the deletion in each candidate.

The phenotype of either N100 (*gal*$E^+T^+K^-$) [3] or MI101 (*Δgal*) cells [8] containing the recombinant plasmids was checked on MacConkey galactose indicator plates [11]. Proteins in crude extracts were analysed by SDS-polyacrylamide gel electrophoresis [10] and galactokinase activities were measured by the standard assay [3]. DNA sequences were determined by the Maxam-Gilbert method [12].

3. RESULTS AND DISCUSSION

The construction described in fig.1 resulted in a set of 50 different plasmid derivatives: in each one, the N-terminal sequence of the *gal*E gene up to the *Bst*EII site was fused to a point either before or in the *gal*K gene (fig.1c). We reasoned that some plasmids would carry in-phase fusions between *gal*E and the coding region of *gal*K and that some of these may code for a hybrid protein that was active. To test this, we examined the effect of galactose upon growth of MI101 cells carrying each of the plasmids: it is known that cell growth is inhibited by galactose when galactokinase activity is expressed in a *gal*E or *gal*T background [5]. Galactose inhibited the growth of all cells carrying deletions that did not enter the *gal*K gene whereas it had no effect when the deletions removed more than 40 bp of the gene. Growth was blocked in some cases when up to 40 bp had been removed from the gene. For further study, we chose the candidate with the longest deletion whose growth was stopped by galactose in MI101 (*Δgal*) cells. This

plasmid, pAA121, conferred a Gal$^+$ phenotype on MacConkey galactose indicator plates, when transferred to the *gal*K host, N100. Further, these cells grew well in minimal media containing galactose, showing that the hybrid *gal*E–*gal*K protein complemented the *gal*K mutation in the host. However, we were unable to measure any significant galactokinase activity in cell extracts. As the assay would have detected at least 1% of the level made by pAA113, we conclude that the activity of the hybrid protein is unstable in cell extracts. Attempts to stabilise the activity by adding glycerol, extra salt or proteinase inhibitors have been unsuccessful to date.

The DNA sequence of the hybrid gene in pAA121 was determined, and the zone covering the N-terminal is shown in fig.2, together with the corresponding sequence of the wild-type *gal*K gene on pAA113. Comparison of the two sequences shows that the fusion was formed from a deletion of the first 31 bp of the gene that removed 11 amino acids. These are replaced by 9 amino acids comprising, as expected, the first 6 amino acids of *gal* up to the *Bst*EII site, and 3 amino acids introduced by the linker. The sequence confirmed that the *gal*E and *gal*K reading frames were in phase.

SDS-polyacrylamide gel analysis of extracts of cells carrying pAA121 (fig.3) showed a clear band corresponding to the hybrid protein (lane f). As controls, we analysed extracts of cells carrying pAA113 (lane a), two plasmids where deletions had not entered *gal*K (lanes b,c) and a plasmid where a deletion removed 700 bp of *gal*K (lane d). To prove that translation of the hybrid protein initiated at the correct position, we made two derivatives of pAA121 by replacing the *Eco*RI–*Bst*EII fragment carrying the *gal* promoter region (fig.1c) with fragments carrying either a mutated *gal*E initiator codon or a defective Shine-Dalgarno sequence [8]. The former change results in an almost total disappearance of the hybrid-protein band (fig.3g) whereas the latter reduces the level by 4-fold (fig.3e). Both changes cause distinctive alterations in the Gal phenotype of N100 host cells on indicator plates. Because this phenotype reflects the rate of initiation of *gal*E translation, it is possible to use pAA121 to select for mutations that alter this. This plasmid may also be useful to study translation initiation in other systems as it allows the replacement of the N-terminal of *gal*K with the N-terminal zone of other genes.

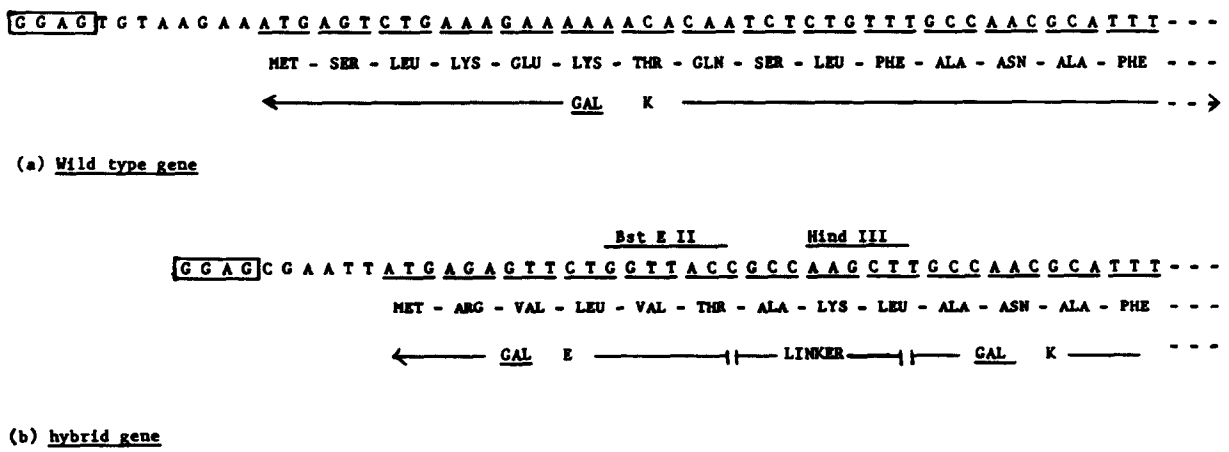

Fig.2. (a) The sequence of the N-terminal zone of the wild-type *gal*K gene. The coding strand is shown together with the translation product [9]. A short stretch of the sequence before the gene is shown and the Shine-Dalgarno sequence is boxed. (b) The sequence of the N-terminal segment of the hybrid gene on pAA121. The coding strand and translation product are shown together with a short stretch of the non-translated leader sequence on which the Shine–Dalgarno sequence is boxed. The sequence is aligned with that in (a), in order to highlight the zones of the protein that derive from *gal*E, the linker and the *gal*K gene. The position of the *Bst*EII and *Hind*III sites are shown.

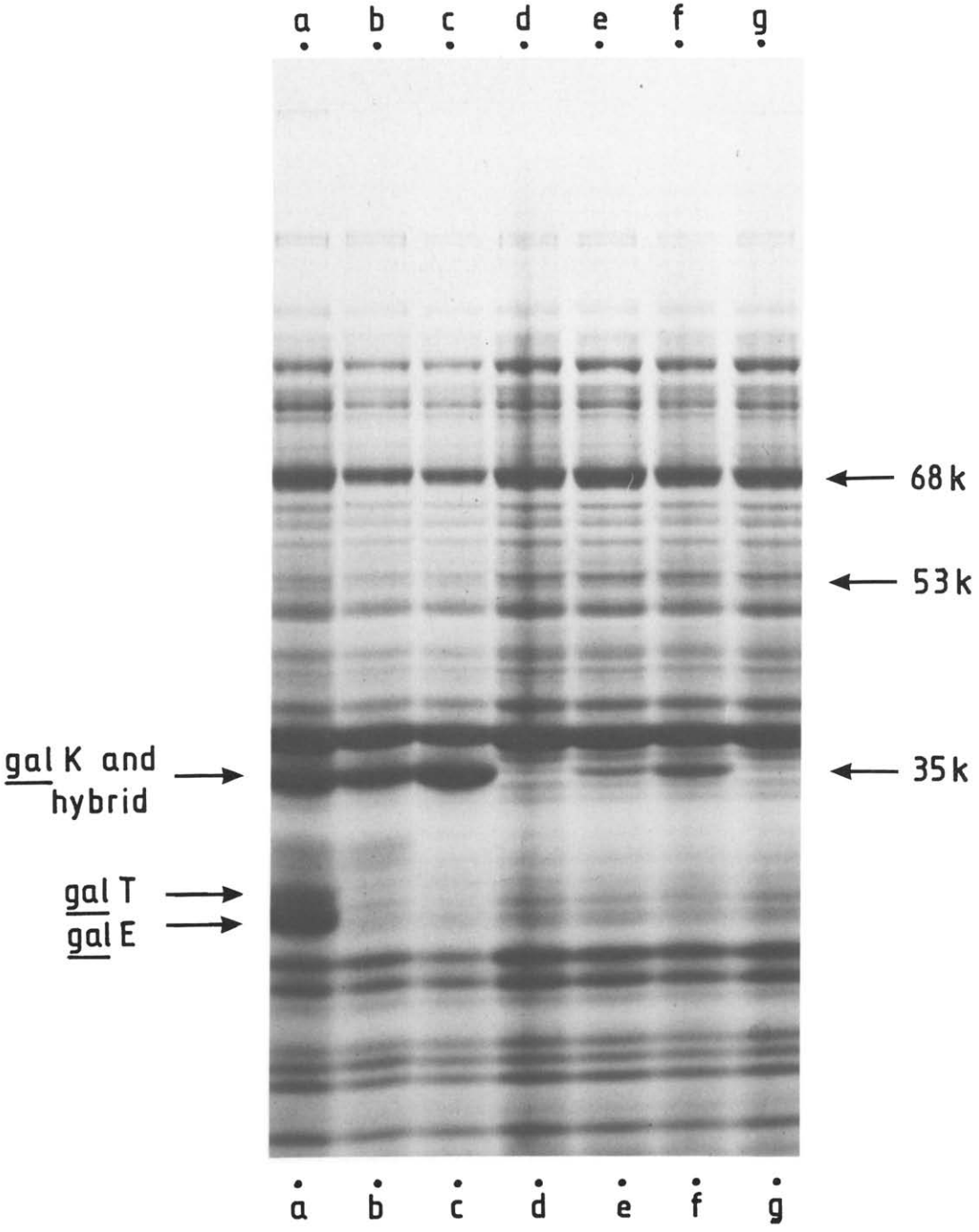

Fig.3. Polyacrylamide gel analysis of extracts of MI101 cells containing various plasmids: (a) pAA113; (b,c) derivatives containing deletions from the *Bst*EII site to a point just before *gal*K; (d) derivative containing a deletion that removes 700 bp of *gal*K; (e) pAA121 carrying a mutation that converts the *gal*E Shine-Dalgarno sequence from 5′GGAG3′ to 5′GAAG3′; (f) pAA121 carrying the wild-type *gal* regulatory region; (g) pAA121 carrying a mutation in the *gal*E initiator codon that converts AUG to AUA. The gel was coloured with Coomassie brilliant blue and was calibrated with standard proteins. The positions of the three *gal* specific bands were deduced from this gel and from previous work [8].

ACKNOWLEDGEMENT

This work was supported by an SERC project grant.

REFERENCES

[1] Shilhavy, T., Berman, M. and Enquist, L. (1983) in: Experiments with Gene Fusions, Cold Spring Harbor Laboratory, New York.
[2] Casadaban, M. and Cohen, S. (1980) J. Mol. Biol. 138, 179–207.
[3] McKenney, K., Shimatake, H., Court, D., Schmeissner, U. and Rosenberg, M. (1981) in: Gene Amplification and Analysis (Chirikjian, J. and Papas, T. eds) vol.2, pp.383–415, Elsevier, Amsterdam, New York.
[4] Brosius, J. (1984) Gene 27, 151–160.
[5] Rosenberg, M., Chepelinsky, A. and McKenney, K. (1983) Science 222, 734–739.
[6] Muller-Hill, B. and Kania, J. (1974) Nature 249, 561–562.
[7] Casadaban, M., Chou, J. and Cohen, S. (1980) J. Bacteriol. 143, 971–980.
[8] Busby, S. and Dreyfus, M. (1983) Gene 21, 121–131.
[9] Schumperli, D., McKenney, K., Sobeiski, D. and Rosenberg, M. (1982) Cell 30, 865–871.
[10] Maniatis, T., Fritsch, E. and Sambrook, J. (1982) in: Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory, New York.
[11] Miller, J. (1972) in: Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, New York.
[12] Maxam, A. and Gilbert, W. (1980) Methods Enzymol. 65 (Grossman, L. and Moldave, K. eds) pp.499–560, Academic Press, New York.