blotting (unpublished results) it was concluded that CV-1 cells infected with SVEHA3 contain about five times more HA mRNA than do cells infected with influenza virus. This excessive quantity of HA reflects in part the fact that cells infected with the recombinant vectors accumulate HA during the virus growth cycle, whereas cells infected with influenza virus shed HA as a normal part of the process of budding progeny virions. However, it also reflects the efficiency with which the HA gene is expressed from a recombinant DNA virus which replicates to high copy number.

Influenza HA therefore joins a number of other foreign genes which have been inserted into late replacement vectors of SV40 and expressed in simian cells[33–36]. Although it is difficult to obtain accurate estimates from the published data, it is clear that the level of expression of these genes is highly variable. Whereas cells infected with a late replacement vector carrying a cDNA copy of the rat preproinsulin gene synthesized $\sim 10^6$ copies of proinsulin per cell[36], those infected with vectors carrying cDNA or genomic copies of β-globin DNA synthesized $>10^7$ copies per cell[33,34]. In only one system, however, is the foreign gene expressed in amounts comparable with influenza HA. Mulligan *et al.*[33] report that simian cells infected with SVGT-RaβG, which carries the cDNA sequences for rabbit β-globin in place of the gene coding for the SV40 capsid protein VP1, synthesize rabbit β-globin polypeptide in quantities nearly equivalent to the amount of VP1 produced by the helper virus. In our system influenza HA can be seen as a major band in extracts of radiolabelled vector-infected cells. We estimate that about 5–10 times more HA than VP1 is produced in cells containing approximately equal numbers of helper and vector genomes.

The late replacement vector contains none of the splice donor or acceptor sites used in the synthesis of late SV40 mRNAs. It therefore seems likely that HA, like several other viral and cellular proteins[37–41], is efficiently translated from an unspliced mRNA. A study of the structure of the HA mRNAs synthesized by SVEHA3 and SVLHA8 will be published elsewhere.

These experiments open the way to use site-directed mutagenesis to dissect those parts of the molecule important in its structure, function and biosynthesis and which have hitherto been unaccessible to genetic analysis.

Furthermore, it is now feasible to express hybrid proteins at the cell surface by inserting foreign nucleotide sequences into the vectors between the HA sequences which code for the hydrophobic signal and the membrane anchoring peptide. Modification of these vectors by removing the anchoring sequences may allow proteins which are normally found only in intracellular locations or attached to the membrane to be secreted from the host cells.

We thank Jackie Bye and Ronni Greene for technical assistance and Ed Harlow for helpful discussions and critical reading of the manuscript. We are grateful to Yasha Gluzman for providing COS-1 cells. This work was facilitated by NATO Collaborative Research Grant RG160.80. The work at Cold Spring Harbor was supported by the NCI.

Received 20 July; accepted 27 August 1981.

1. Tomita, M. & Marchesi, V. T. *Proc. natn. Acad. Sci. U.S.A.* **72,** 2964–2968 (1975).
2. Orr, H. T. *et al. Proc. natn. Acad. Sci. U.S.A.* **76,** 4395–4399 (1979).
3. Campbell, D. G. *et al. Biochem. J.* **195,** 15–30 (1981).
4. Coligan, J. E. *et al. Nature* **291,** 35–39 (1981).
5. Ward, C. W. & Dopheide, T. A. A. *Virology* **103,** 37–53 (1980).
6. Dopheide, T. A. A. & Ward, C. W. *J. gen. Virol.* **50,** 329–335 (1980).
7. Porter, A. G. *et al. Nature* **282,** 471–477 (1979).
8. Gething, M. J. *et al. Nature* **287,** 301–306 (1980).
9. Both, G. W. & Sleigh, M. J. *Nucleic Acids Res.* **8,** 2561–2575 (1980).
10. Min Jou, M. *et al. Cell* **19,** 683–696 (1980).
11. Verhoeyen, M. *et al. Nature* **286,** 771–776 (1980).
12. Winter, G., Fields, S. & Brownlee, G. G. *Nature* **292,** 72–75 (1981).
13. Hiti, A. L., Davis, A. R. & Nayak, D. P. *Virology* **111,** 113–124 (1981).
14. Fields, S., Winter, G. & Brownlee, G. G. *Nature* **290,** 213–217 (1981).
15. Garoff, H. *et al. Nature* **288,** 236–241 (1980).
16. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289,** 366–373 (1981).
17. Wiley, D. C., Wilson, I. A. & Skehel, J. J. *Nature* **289,** 373–378 (1981).
18. Davis, B. D. & Tai, P. C. *Nature* **283,** 433–437 (1980).
19. Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 93–97 (1975).
20. Ploegh, H. L., Orr, H. T. & Strominger, J. L. *Cell* **24,** 287–299 (1981).
21. *Molecular Biology of Tumour Viruses* 2nd edn, Pt 2 (ed. Tooze, J.) (Cold Spring Harbor Laboratory, New York, 1980).
22. Ghosh, P. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **78,** 100–104 (1981).
23. Pipas, J. M. *et al. Cold Spring Harb. Symp. quant. Biol.* **44,** 285–291 (1980).
24. McCutchan, J. H. & Pagano, J. S. *J. natn. Cancer Inst.* **41,** 351–357 (1968).
25. Emtage, J. S. *et al. Nature* **283,** 171–174 (1980).
26. Elder, K. T. *et al. Virology* **95,** 343–350 (1979).
27. Laver, W. G. *Virology* **45,** 275–288 (1971).
28. Horsfall, F. L. & Tamm, I. J. *Immun.* **70,** 253–259 (1953).
29. Gluzman, Y. *Cell* **23,** 175–182 (1981).
30. McCormick, F. & Harlow, E. *J. Virol.* **34,** 213–224 (1980).
31. Alwine, J. C., Reed, S. I. & Stark, G. R. *J. Virol.* **24,** 22–31 (1977).
32. Hirt, B. J. *molec. Biol.* **26,** 365–369 (1967).
33. Mulligan, R. C., Howard, H. & Berg, P. *Nature* **277,** 108–114 (1979).
34. Hamer, D. H. & Leder, P. *Nature* **281,** 35–40 (1979).
35. Hamer, D. H. *et al. Cell* **17,** 725–735 (1979).
36. Gruss, P. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **78,** 133–137 (1981).
37. Schaffner, W. *et al. Cell* **14,** 655–671 (1978).
38. Nagata, S., Mantei, N. & Weissman, C. *Nature* **287,** 401–408 (1980).
39. Alestrom, P. *et al. Cell* **19,** 671–681 (1980).
40. Ghosh, P. K. *et al. J. molec. Biol.* **126,** 813–848 (1978).
41. Treisman, R. *et al. Nature* (in the press).
42. Fiers, W. *et al. Nature* **273,** 113–120 (1978).
43. Heyneker, H. L. *et al. Nature* **263,** 748–752 (1976).
44. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
45. Luskey, M. & Botchan, M. *Nature* (in the press).
46. Weaver, R. F. & Weissman, C. *Nucleic Acids Res.* **7,** 1175–1193 (1979).
47. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).
48. Harlow, E., Pim, D. C. & Crawford, L. V. *J. Virol.* **37,** 564–573 (1981).
49. Hay, A. J. *Virology* **60,** 398–418 (1974).
50. Laver, W. G. *J. molec. Biol.* **9,** 109–124 (1964).
51. Fraker, P. J. & Speck, J. C. *Biochem. biophys. Res. Commun.* **80,** 849–857 (1978).
52. Favaloro, J., Triesman, R. & Kamen, R. *Meth. Enzym.* **65,** 718–749 (1980).
53. Ey, P. L., Prowse, S. J. & Jenkin, C. R. *Immunochemistry* **15,** 429–436 (1978).
54. Gray, H. B. *et al. Nucleic Acids Res.* **2,** 1459–1492 (1975).
55. Choppin, P. W., Lazarowitz, S. G. & Goldberg, A. R. in *Negative Strand Viruses*, vol. 1 (eds Mahy, B. W. J. & Barry, R. D.) 105–119 (Academic, London, 1975).

Insertion of DNA activates the cryptic *bgl* operon in *E. coli* K12

Ann E. Reynolds, Jeffrey Felton* & Andrew Wright

Department of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, Massachusetts 02111, USA

Spontaneous mutations which activate the cryptic bgl *operon of* Escherichia coli *K12 are caused by insertion of DNA at a site,* bglR, *within the operon. Two insertion elements, IS*1 *and IS*5, *have been observed to effect this activation. Once the activating insertion has occurred the operon is inducible by β-glucosides in a cyclic AMP-dependent manner.*

TRANSPOSABLE DNA elements in bacteria are known to be involved in site-specific recombination events such as inversion, deletion and transposition[1]. The insertion of such elements into a gene usually results in inactivation of that gene and of distal genes in the same transcriptional unit. However, the insertion of three transposable elements, IS*2*, Tn*A* and Tn*5* (refs 2, 3 and 4 respectively), can sometimes lead to activation of a downstream gene or operon due to the presence of promoters within these transposons. Expression from these promoters is unregulated and requires a unique orientation or location of the transposon relative to the gene being expressed. We demonstrate here that activation of the cryptic *bgl* operon in *E. coli* K12 occurs

* Present address: Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.

0028-0836/81/430625—05$01.00

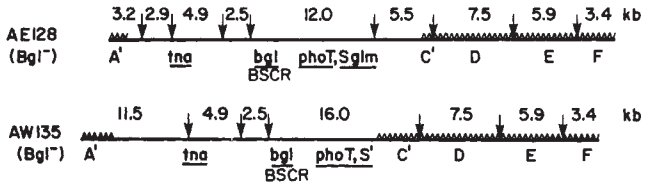

Fig. 1 λ specialized transducing phages carrying the *bgl* operon. The specialized transducing phage λAE128 was isolated by a two-step procedure starting with a lysogen (JF225) containing λY199 (λcI857 xis(Am) *S*7b515b519) integrated at a secondary attachment site in the *phoS* gene of *E. coli* K12. Induction of JF225 gave rise to a specialized transducing phage carrying the region from *phoS* through *tnaA*, identified by its ability to transduce a *tnaA*⁻ cell to *tnaA*⁺. A lysogen (JF369) carrying this transducing phage, which was integrated by homology in the *tnaA* region of the chromosome, was induced giving rise to specialized transducing phages carrying the chromosomal region from *tnaA* to *glmS*. λAE128 is one such phage, identified by its ability to transduce a *glmS*⁻ cell to *glmS*⁺ (ability to synthesize glucosamine). λAE128 is a defective phage which carries the chromosomal alleles *glmS*⁺, *tnaA*⁻ and *bglR*⁰. λAW135 was isolated by inducing a *phoS*::λY199 lysogen (AE120), infecting a *tnaA*⁻ strain with the resulting lysate and selecting *tnaA*⁺ transductants. One of these transductants, AW135, was used as the source of the *tnaA*⁺ *bglR*⁰ transducing phage λAW135. λAE128 has an *Eco*RI restriction site in the segment of bacterial DNA between A' and *tna* which is absent in the equivalent segment of λAW135 bacterial DNA. This restriction site is characteristic of the particular strain from which λAE128 was derived. —— Indicates bacterial DNA, ~~ indicates λ DNA, ↓ indicates the *Eco*RI restriction sites.

spontaneously by insertion of either IS*1* or IS*5* into the operon. In contrast to the above examples of activation, expression of this 'insertion-activated' operon is subject to substrate-specific regulation. The mechanism which we describe for activation of the *bgl* operon represents a novel system of genetic regulation by transposition of insertion sequences.

The *bgl* operon, which is located near minute 83 on the *E. coli* K12 genetic map, contains three structural genes required for the catabolism of aromatic β-glucosides, such as salicin and arbutin[5]. The operon is considered to be cryptic because these genes cannot be expressed, or induced, in wild-type *E. coli* K12. Wild-type bacteria are thus unable to ferment salicin (Sal⁻) or arbutin (Arb⁻). However, Bgl⁺ (Sal⁺, Arb⁺) mutants can arise spontaneously, and in these mutants the operon is inducible by various β-glucosides. The genes of the *bgl* operon are *bglB*, which encodes a hydrolytic enzyme, phospho-β-glucosidase B; *bglS*, whose product is a regulator of the *bgl* operon; and *bglC*, which encodes a specific transport activity[6]. In addition to these three structural genes there exists a site, *bglR*, which is required for their expression, and where the spontaneous mutations allowing expression of the *bgl* operon occur[5]. We have termed the mutations giving rise to the Bgl⁺ phenotype, *bglR*, and have designated the wild-type allele *bglR*⁰, because it gives a Bgl⁻ phenotype. (The nomenclature used here for the *bglR* site differs from that used by Prasad and Schaefler[5] in the following ways. We have used *bglR*⁰ in place of *bglR*⁺ to indicate the cryptic nature of the wild-type site, and *bglR* instead of *bglR*⁻ to indicate the mutant site which allows expression of the *bgl* operon.)

Figure 1 shows the organization of the genes of the *bgl* operon. This gene order, which differs from that originally proposed by Prasad and Schaefler[5], was determined by a physical and genetic analysis of the *bgl* operon cloned in plasmid pBR322 (in preparation). Merodiploid analysis has shown that *bglR* is dominant to *bglR*⁰ and that expression of genes *B*, *S* and *C* requires a *bglR* allele *cis* to the structural genes[5]. The S product is *trans*-acting and according to the model proposed by Prasad and Schaefler[5] is a positive regulator required for expression of genes *bglB* and *bglC*. The location of the *bglR* mutation and its *cis* action on the expression of the operon strongly suggest that transcription of the *bgl* operon is initiated near *bglR*.

The requirements for the *bglR* mutation, with its *cis*-acting effects, are so specific that it would be expected to occur only at very low frequency. In fact, it occurs at frequencies as high as 10^{-5} in some *E. coli* K12 strains, suggesting that a special mechanism may exist for the spontaneous *bglR*⁰ to *bglR* mutation. We considered that DNA rearrangements might be involved and tested this hypothesis by comparing the DNA structure of operons containing *bglR*⁰ and *bglR* alleles.

Specialized transducing phages carry the *bgl* operon

To examine the structure of the *bgl* operon, we isolated the λdBgl⁻ specialized transducing phages shown in Fig. 1. These phages were generated by induction of lysogens containing λ prophages in secondary attachment sites near the *bgl* operon. Two Bgl⁻ phages were used in this work; λAE128, which carries a segment of bacterial chromosome from *tnaA* through *glmS*, and λAW135, which carries DNA from *tnaA* to *phoS*. The transducing phages λAE129 and λAE130 are independent spontaneous Bgl⁺ derivatives of λAE128, and λAW138 is a spontaneous Bgl⁺ derivative of λAW135.

We assayed phospho-β-glucosidase B activity in cells lysogenic for the series of Bgl⁻ and Bgl⁺ transducing phages to show that the phage-carried operon is subject to the same regulation as the operon in its normal chromosomal location. The results (see Table 1) indicate that the Bgl⁻ lysogens AE128 and AW135 showed no phospho-β-glucosidase B activity either in the absence or presence of inducer. In contrast, the Bgl⁺ lysogens AE129, AE130 and AW138 did show phospho-β-glucosidase B activity, but only when inducer was present. The Bgl⁻ and Bgl⁺ lysogens are therefore regulated in the same way as the non-lysogenic haploid Bgl⁻ strain RVA and its Bgl⁺ derivative. The Bgl⁺ lysogens exhibit a range of enzyme levels which correlates with the intensity of their colour development on MacConkey salicin indicator plates.

After activation by a *bglR* mutation, expression of the *bgl* operon, in addition to being inducible by β-glucosides, is dependent on cyclic AMP. We showed this by constructing a *bglR* strain (Δ*cya*) which carries a deletion of the gene for adenylate cyclase, *cya*. This strain is Lac⁻ Mal⁻ Bgl⁻ on indicator plates, unlike its isogenic *cya*⁺ parent which is Lac⁺ Mal⁺ Bgl⁺. The defect in the Δ*cya* strain can be corrected by the addition of

Table 1 Specific activity of phospho-β-glucosidase B enzyme in Bgl⁻ and Bgl⁺ strains

Strain	*bgl* genotype	Phenotype	Phoso-β-glucosidase B activity Without inducer	With inducer	Induced activity
RVA	*bglR*⁰	Bgl⁻	0.000	0.007	0.007
RVA (Bgl⁺)	*bglR*	Bgl⁺	−0.013	0.66	0.673
AW135	λ*bglR*⁰/*bglR*⁰	Bgl⁻	0.02	0.035	0.015
AW138	λ*bglR*/*bglR*⁰	Bgl⁺	0.068	0.458	0.390
AE128	λ*bglR*⁰/*bglR*⁰	Bgl⁻	0.013	0.014	0.001
AE129	λ*bglR*/*bglR*⁰	Bgl⁺	0.034	0.222	0.188
AE130	λ*bglR*/*bglR*⁰	Bgl⁺	−0.007	0.013	0.020
AE141	λ*bglR*/*bglR*⁰	Bgl⁺	0.017	0.039	0.022
KO618	*bglR*	Bgl⁺	−0.047	0.131	0.178

Cells were grown to mid-log phase in M9 medium containing succinate (0.4%) as a carbon source, casamino acids (0.2%), thiamine (10 μg ml⁻¹), tryptophan (50 μg ml⁻¹) and with or without β-methylglucoside (5×10^{-3} M) as inducer. Cells were washed with 0.8% saline and resuspended in saline to one-tenth of the original volume. Phospho-β-glucosidase B activity was assayed as follows[20]. 0.1 ml of 0.14 M salicin and 0.1 ml of cell suspension were mixed and incubated for 30 min at 37 °C. The reaction was stopped by addition of 0.5 ml of 2M Na_2CO_3. To each sample, at room temperature, 0.5 ml of a 0.6% solution of 4-amino-antipyrine was added, followed 15 min later by 0.5 ml of 4% $K_3Fe(CN)_6$. After 5 min at room temperature the optical density (A) at 509 nm was measured. This measurement was corrected for the contribution made by turbidity. Specific activity was determined by dividing the corrected A_{509} by the optical density of the cell culture measured at 600 nm. The values in the table are given relative to phospho-β-glucosidase B activity in the wild-type haploid strain (RVA) when grown in the absence of inducer. The control level of activity for lysogens AE129, AE130 and AE141 is that given by the parent Bgl⁻ lysogen AE128. The Bgl⁻ lysogen, AW135, is the analogous control for AW138. AW135, AW138, AE128, AE129 and AE130 are lysogens carrying the λBgl⁻ and λBgl⁺ specialized transducing phages as described in the text. AE141 is lysogenic for λBgl⁺ phage derived by mutagenesis of the λBgl⁻ phage present in AE128. KO618 is a Bgl⁺ strain derived by mutagenesis of the Bgl⁻ haploid strain RVA. Mutagenesis was carried out by placing filter disks, spotted with either ethylmethanesulphonate or nitrosoguanidine, on MacConkey salicin agar which had been spread with the Bgl⁻ strain. Red papillae (Bgl⁺ mutants) which appeared were isolated and purified for study.

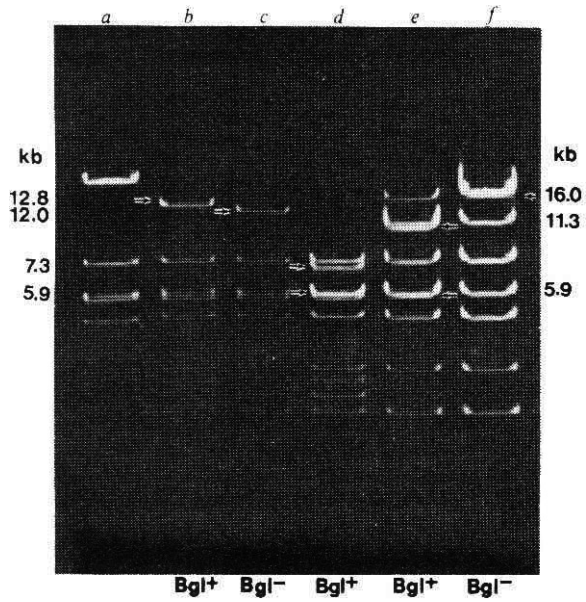

Fig. 2 *Eco*RI restriction fragment patterns of Bgl⁻ and Bgl⁺ specialized transducing phages. Phages were obtained by heat induction of the appropriate lysogen and purified in CsCl gradients. DNA was extracted from the phages with phenol and samples (0.75 μg) were digested with *Eco*RI restriction endonuclease. DNA fragments were separated by electrophoresis on 0.9% agarose using a buffer containing 50 mM sodium acetate, 10 mM EDTA and 0.4 M Tris pH 7.9 for 2 h at 120 V. An *Eco*RI digest of λcI857 S7(Am) DNA was used as a size standard[21]. *a*, λcI857S7; *b*, λAE130; *c*, λAE128; *d*, λAE129; *e*, λAW138; *f*, λAW135.

1 mM cyclic AMP to the indicator plates, thereby restoring the Lac⁺ Mal⁺ Bgl⁺ phenotype.

Restriction enzyme analysis of Bgl⁻ and Bgl⁺ transducing phages

Figure 2 shows the *Eco*RI restriction fragment patterns of the λ specialized transducing phages λAE128, λAW135 and their respective isogenic Bgl⁺ derivatives. Cloning evidence (A.E.R., unpublished observations) has revealed that the largest *Eco*RI fragment (12.0 kilobases [kb]) of λAE128 contains the entire *bgl* operon and Southern hybridization has shown that the fragment consists only of chromosomal DNA. In the case of λAW135, the largest fragment (16.0 kb) contains almost all of the *Eco*RI fragment carrying the *bgl* operon and part of λ*Eco*RI fragment C (see Fig. 1). The order of restriction fragments in the transducing phages is based on mapping data of the *oriC* region of the *E. coli* chromosome[7].

The *bglR*⁰-containing *Eco*RI fragments of λAE128 and λAW135 are altered by spontaneous mutation to *bglR*. The uppermost fragment of λAE130 (Fig. 2) is slightly larger than that of the parent *bglR*⁰ phage λAE128, indicating that additional DNA is present in the *bglR* derivative. λAE129 and λAW138 lack the uppermost fragment present in their respective parental phages but contain two new fragments, 7.3 kb and 5.9 kb in the case of λAE129, and 11.3 kb and 5.9 kb in the case of λAW138. In each case the 5.9-kb fragment coincides with *Eco*RI fragment E of phage λ, as the band has the intensity of a doublet. These results with *Eco*RI and results from analyses using *Hin*dIII and *Pst*I (data not shown) indicate that mutation from Bgl⁻ to Bgl⁺ is correlated in each case with insertion of DNA.

To measure more accurately the sizes of the insertions, we compared the *Hin*cII + *Hin*dIII restriction patterns of the Bgl⁻ and Bgl⁺ transducing phages. As shown in Fig. 3, a 910-base pair (bp) fragment which is present in both Bgl transducing phages is missing in all the Bgl⁺ derivatives, indicating that the insertion affecting the *bglR* locus has occurred somewhere within the same 910-bp segment of DNA in each spontaneous Bgl⁺ mutant. In λAE130 a new, larger fragment forms a doublet with the 1.70-kb fragment already present in the parent phage. This fragment was clearly resolved in a polyacrylamide-agarose gel system (result not shown). Thus the insertion in λAE130 is 790 bp. A new fragment of 2.15 kb is seen in both λAE129 and λAW138, indicating an insertion of 1,240 bp.

Thus two different DNA insertion events are correlated with the spontaneous mutations from Bgl⁻ to Bgl⁺. Eight independent spontaneous Bgl⁺ mutations have been analysed, using *bgl* phages, as described above; four were found to be due to the smaller insertion (790 bp), the other four to the larger insertion (1,240 bp). Each insertion was found to occur in Bgl⁺ derivatives of either of the Bgl⁻ parent phages, λAE128 and λAW135. Because the wild-type *bgl* operon present on λAE128 originated from a different *E. coli* K12 strain line than that giving rise to λAW135, we believe that activation of the *bgl* operon by insertion of DNA is a general phenomenon.

To determine whether the DNA insertions present in the spontaneous Bgl⁺ derivatives are related to known insertion sequences, a series of DNA probes containing IS*1*, IS*2*, IS*3* and IS*5* were used for Southern hybridization analysis of restriction fragments from Bgl⁻ and Bgl⁺ phages. Only those containing IS*1* and IS*5* hybridized specifically to *bgl* operon DNA from λBgl⁺ transducing phages.

Bacteriophage P1 DNA, which contains IS*1* (ref. 8), hybridized specifically to the *bgl*-containing *Eco*RI fragment of λAE130 DNA (Fig. 4). It failed to hybridize to DNA from either the parent Bgl⁻ transducing phage, λAE128, its other Bgl⁺ derivative, λAE129, or the pair of phages λAW135 (Bgl⁻) and λAW138 (Bgl⁺), but did hybridize with the *lac*-containing *Eco*RI fragment of λ*plac*3-MS348 which contains IS*1*[9]. These results indicate that the insertion in the DNA of the Bgl⁺ phage λAE130 is homologous to IS*1*. Similar results indicating that the insertion in λAE130 is IS*1* were obtained when λ*plac*3-MS348 was used as a probe. The size of the insertion in λAE130 (determined above) and the presence of *Pst*I, *Pvu*II and *Bal*I sites (A.E.R., unpublished observations) within the insertion[10] is consistent with this conclusion.

Plasmid pGM1 was used as a probe for IS*5*. This plasmid[11] is a pBR322 derivative which contains about 1,150 bp of IS*5*, a segment of bacteriophage Mu DNA and a small portion of IS*2* (see Fig. 4 legend). Plasmid pGM1 hybridized with one of the two *bglR*-specific *Eco*RI subfragments of λAE129 (7.3 kb) and λAW138 (11.3 kb). It did not hybridize to the *bgl* fragment of

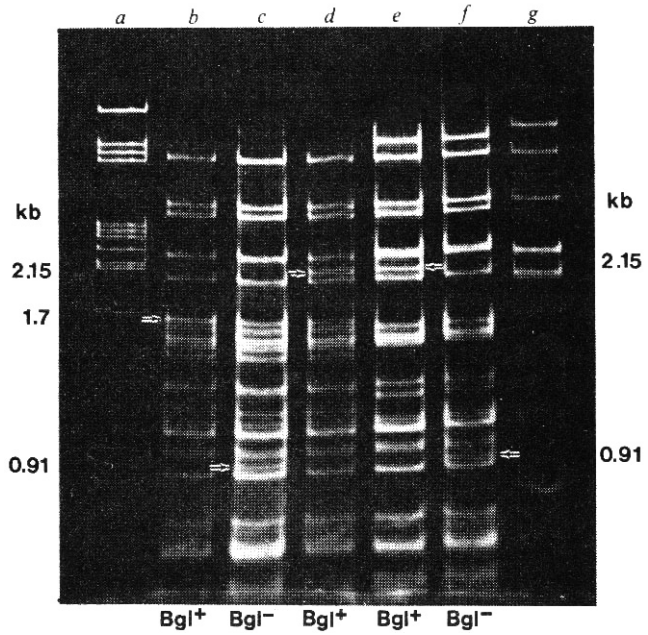

Fig. 3 *Hin*cII-*Hin*dIII restriction fragment patterns of Bgl⁻ and Bgl⁺ specialized transducing phages. Phage DNA was digested with *Hin*cII-*Hin*dIII restriction endonuclease mixture (New England Biolabs). The DNA fragments were separated by electrophoresis on 1.4% agarose at 120 V for 2 h. The size standard is a *Hpa*I digest of phage T7 DNA[22]. *a*, T7, *Hpa*I digest; *b*, λAE130; *c*, λAE128; *d*, λAE129; *e*, λAW138; *f*, λAW135; *g*, λcI857S7, *Hin*cII–*Hin*dIII digest.

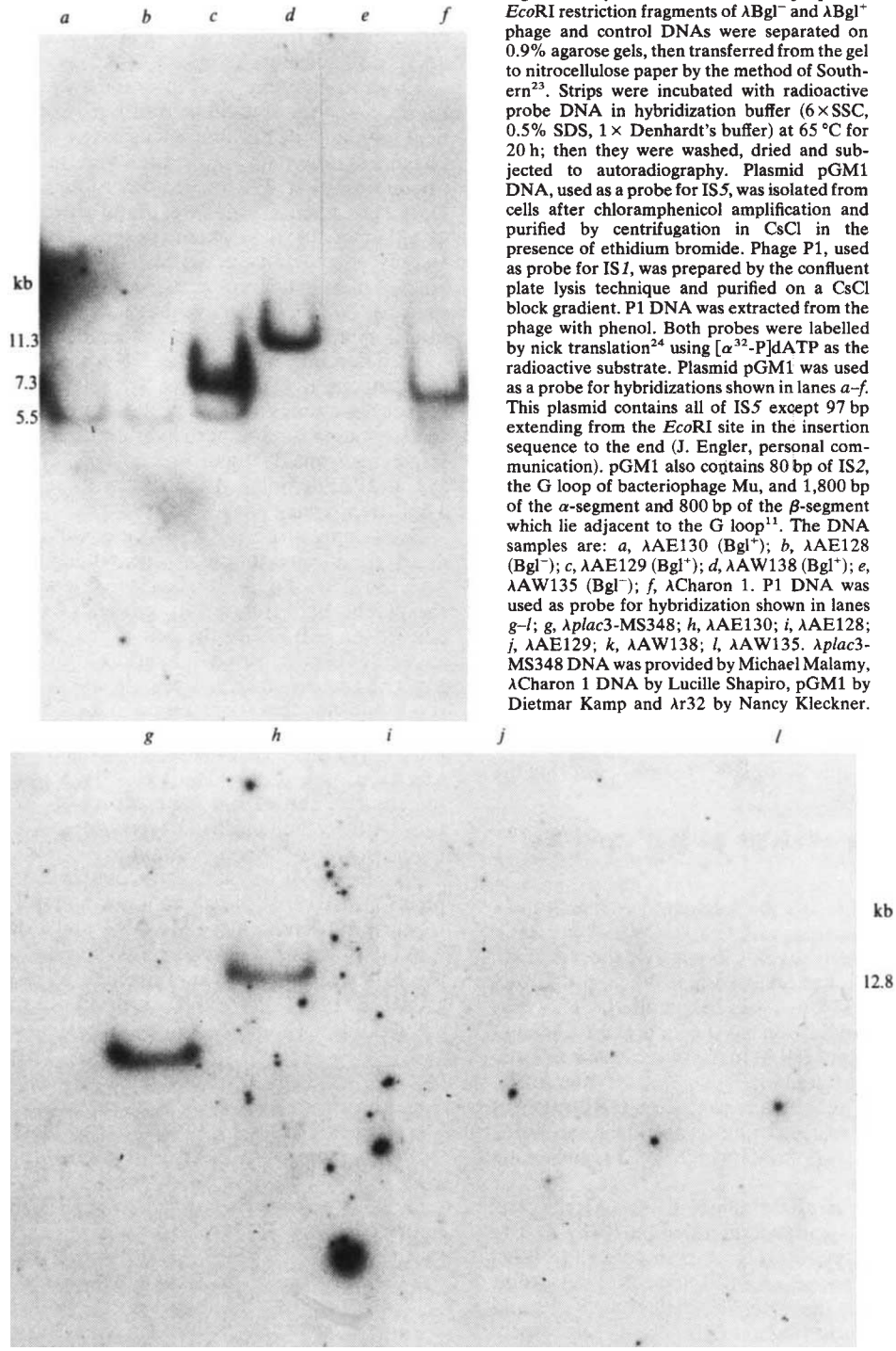

Fig. 4 Analysis of insertions in the *bgl* operon. *Eco*RI restriction fragments of λBgl⁻ and λBgl⁺ phage and control DNAs were separated on 0.9% agarose gels, then transferred from the gel to nitrocellulose paper by the method of Southern[23]. Strips were incubated with radioactive probe DNA in hybridization buffer (6×SSC, 0.5% SDS, 1× Denhardt's buffer) at 65 °C for 20 h; then they were washed, dried and subjected to autoradiography. Plasmid pGM1 DNA, used as a probe for IS*5*, was isolated from cells after chloramphenicol amplification and purified by centrifugation in CsCl in the presence of ethidium bromide. Phage P1, used as probe for IS*1*, was prepared by the confluent plate lysis technique and purified on a CsCl block gradient. P1 DNA was extracted from the phage with phenol. Both probes were labelled by nick translation[24] using [α^{32}-P]dATP as the radioactive substrate. Plasmid pGM1 was used as a probe for hybridizations shown in lanes *a–f*. This plasmid contains all of IS*5* except 97 bp extending from the *Eco*RI site in the insertion sequence to the end (J. Engler, personal communication). pGM1 also contains 80 bp of IS*2*, the G loop of bacteriophage Mu, and 1,800 bp of the α-segment and 800 bp of the β-segment which lie adjacent to the G loop[11]. The DNA samples are: *a*, λAE130 (Bgl⁺); *b*, λAE128 (Bgl⁻); *c*, λAE129 (Bgl⁺); *d*, λAW138 (Bgl⁺); *e*, λAW135 (Bgl⁻); *f*, λCharon 1. P1 DNA was used as probe for hybridization shown in lanes *g–l*; *g*, λ*plac*3-MS348; *h*, λAE130; *i*, λAE128; *j*, λAE129; *k*, λAW138; *l*, λAW135. λ*plac*3-MS348 DNA was provided by Michael Malamy, λCharon 1 DNA by Lucille Shapiro, pGM1 by Dietmar Kamp and λr32 by Nancy Kleckner.

the parent Bgl⁻ phages λAE128 and λAW135 or to that of the Bgl⁺ phage λAE130. pGM1 hybridized to λCharon 1 DNA which contains only IS*5* in common with the plasmid[12]. The *bglR*-specific fragments of λAE129 and λAW138 which hybridize to pGM1 did not hybridize with Mu DNA, λr32 (IS*2*)[13] DNA or pBR322 DNA. Thus the insertion in the *bgl* region of λAE129 and λAW138 is homologous to IS*5*. The size of the DNA insertion in these two phages and the presence of an *Eco*RI site within the insertion (previous section) lead us to conclude that this insertion is IS*5* in both cases. From this hybridization data and the restriction analysis of the phages we conclude that the orientation of IS*5* in *bglR* is the same in both cases. Within the limits of resolution of the gel system used, the point of insertion is also the same, being 5.8 kb from the *Eco*RI site which is indicated at the immediate left end of the *bgl* operon in Fig. 1.

pGM1 also hybridized to another *Eco*RI fragment present in the family of phages λAE128 (Bgl⁻), λAE129 (Bgl⁺) and λAE130 (Bgl⁺). This 5.5-kb fragment lies in the *glmS–uncC,D* region of the chromosome[7] which is not present in λAW135 or λAW138. The presence of IS*5* or a portion of IS*5* in this 5.5-kb fragment is not specific to *bgl*, is well separated from the *bgl* operon and does not seem to be related to expression of the *bgl* operon.

Mutagen-induced *bgl*⁺ strains

The mutagens ethylmethanesulphonate (EMS) and nitrosoguanidine (NTG) increase the frequency with which Bgl⁺ papillae appear in a lawn of a Bgl⁻ strain on MacConkey salicin indicator plates. As mutagens are not known to increase the frequency of transposition of insertion sequences, we wished to determine whether mutagen-induced Bgl⁺ mutations were due to insertions and whether they showed the same regulatory

properties and dominance relationships as spontaneous Bgl^+ mutations. We therefore isolated EMS- and NTG-induced Bgl^+ mutants in both a haploid ($bglR^0$) and merodiploid ($bglR^0/\lambda bglR^0$) background. Twelve Bgl^+ mutants were isolated from the Bgl^- lysogen AE128. Two of the mutations which led to the Bgl^+ phenotype were located on the specialized transducing phage genome. The *Hin*cII–*Hin*dIII and *Eco*RI restriction patterns of these phages were identical to those of the Bgl^- parent phage. Thus neither of these phages contained the IS*1* or IS*5* insertions seen in the spontaneous Bgl^+ mutants. In 10 of the 12 Bgl^+ mutants isolated, the mutations were present on the chromosome and were dominant to $bglR^0$ on the phages.

The haploid strain RVA (Bgl^-) was mutagenized with EMS or NTG and Bgl^+ mutants were selected. In 12 such mutants the *bgl* mutation was P1-cotransducible with *tnaA* at a frequency of 65–90%, which is characteristic of spontaneous *bglR* mutations. When a Bgl^- phage (λAW135) was introduced into strains carrying these *bgl* mutations, the resulting lysogens showed a Bgl^+ phenotype in all cases, indicating dominance of the mutagen-induced *bgl* mutations over the wild type.

The mutagen-induced *bgl* mutations thus appear very similar to spontaneous *bglR* mutations in that they show the same dominance properties and the operons which they activate are inducible by β-glucosides (Table 1). However they are distinctly different from the spontaneous mutations because they show no evidence of an insertion in the *bgl* region.

Discussion

Spontaneous mutations which activate the cryptic *bgl* operon in *E. coli* K12 occur at a specific region within the *bglR* operon[5]. We have demonstrated here that these mutations are due to insertion of IS*1* or IS*5* into a specific region in the operon. Operon expression in the mutants is inducible, dependent on cyclic AMP and requires the presence of the insertion in *cis*. We have considered two different models to explain how insertion of IS*1* or IS*5* could effect the *cis* activation of this operon. In the first model, the *bgl* operon lacks a functional promoter for *bglC*, *bglS* and *bglB*; insertion of DNA into the *bglR* region would either activate a pre-existing promoter or provide or contribute to promoter structure. Because the properties of other insertion mutations produced by IS*1* and IS*5* indicate that these elements do not show promoter activity[1], it seems unlikely that they could provide a complete promoter sequence for transcription of the *bgl* operon. In the second model, the operon contains an operator site which is disrupted by the insertion sequence. The existence of an operator sequence would imply the existence of a repressor of the operon. Such a repressor would be refractory to induction by β-glucosides. Defez and DeFelice[14] have recently described mutations at a locus, *bglY*, which are recessive and which lead to a Bgl^+ phenotype. They suggested that *bglY*, which maps near *trp*, may code for a repressor of the *bgl* operon. However, in our strains, analysis of many spontaneous and mutagen-induced Bgl^+ derivatives has failed to provide any evidence for such a repressor; all Bgl^+ derivatives analysed contained dominant mutations located within the *bgl* operon. The Bgl^+ phenotype conferred by *bglY* mutations may therefore be strain specific. Further work will be necessary to determine the role of the *bglY* product in *bgl* operon expression.

Analysis of two independent IS*5* insertions and one IS*1* insertion indicated that all three are located within 50–100 nucleotides of another. The insertions probably do not occur at the same nucleotide because spontaneous Bgl^+ mutants vary in their levels of operon expression. The small size of the insertion target and the relatively high frequency at which the spontaneous Bgl^- to Bgl^+ mutation occurs leads us to conclude that *bglR* is a preferred site for insertion of both IS*1* and IS*5*.

All the spontaneous mutations that we have analysed have been due to insertion of DNA. However, the frequency of mutation to Bgl^+ can be stimulated by treatment with chemical mutagens. These mutations have similar biological properties to the spontaneous mutations but do not seem to be due to insertion of DNA. Although it is possible that the mutagen-induced mutations activate the operon by a mechanism similar to that used by the insertions, this may not be the case.

Activation of gene expression by site-specific recombination can occur by inversion, as in phase variation in *Salmonella*[15]; by substitution, seen in control of mating type in yeast[16]; by deletion, as occurs during differentiation of immunoglobulin-producing cells in the mouse[17]; and by insertion, such as in the activation of a cellular *onc* gene in virus-induced lymphoid leukosis[18]. All these processes appear to be part of specific systems of biological control; however, there are many examples of alterations of gene expression by non-homologous recombination which do not appear to occur by specific regulatory mechanisms. Examples include inactivation of genes by insertion of transposable genetic elements and nonspecific activation of a gene by integration of an insertion sequence carrying a promoter (for example, IS*2*)[2]. The existence of a preferred site for integration of two IS sequences into the *bgl* operon, the activating effect of such an integration event and the subsequent dual regulation of the activated operon suggest that this activation may be part of a specific system of biological control which confers some evolutionary advantage on cells possessing it.

If this system is used to turn the *bgl* operon of *E. coli* on and off repeatedly, the reverse mutation, *bglR* to $bglR^0$, must also occur and in fact it has been observed in a particular mutant background which has an increased frequency of precise excision of transposons such as Tn*5* and Tn*10* (J. Hopkins and M. Syvanen, personal communication). One possible reason why the type of physiological regulation seen in operons such as the lactose operon might not be adequate for the *bgl* operon is that many β-glucosides are produced in nature, some of which are toxic, for example, the cyanogenic β-glucosides[19]. Such toxic β-glucosides might be inducers of the *bgl* operon and inhibitory to cells which metabolize them, but not to Bgl^- cells. Mutation from *bglR* to $bglR^0$ would enable the cell to survive exposure to toxic β-glucosides while conserving the structural genes of the operon in a form which could be easily reactivated by a *bglR* mutation at a later time.

Activation of cryptic genes by insertion mutations may occur in other systems as well and may be a widespread method of control. This method of mutational genetic regulation may be a specific example of the general class of regulation by recombination.

We thank Dr T. Silhavy for his important contribution to the early stages of this work, S. Michaelis, K. O'Day and W. O'Neil for experimental help, F. Ollington for technical advice, and J. Beckwith, M. Malamy and A. L. Sonenshein for critical reading of the manuscript. A.E.R. was supported by a NSF predoctoral fellowship, J.F. by a NIH training grant fellowship and the research by a NIH grant.

Received 14 May; accepted 26 August 1981.

1. Kleckner, N. *Cell* **11**, 11–23 (1977).
2. Saedler, H., Reif, H. J., Hu, S. & Davidson, N. *Molec. gen. Genet.* **132**, 265–289 (1974).
3. Heffron, F., McCarthy, B. J., Ohtsubo, H. & Ohtsubo, E. *Cell* **18**, 1153–1163 (1979).
4. BaD. E., Weiss, A. & Crossland, L. *J. Bact.* **142**, 439–446 (1980).
5. Prasad, I. & Schaefler, S. *J. Bact.* **120**, 638–650 (1974).
6. Fox, F. & Wilson, G. *Proc. natn. Acad. Sci. U.S.A.* **59**, 988–995 (1968).
7. von Meyenburg, K., Hansen, F. G., Nielsen, L. D. & Jorgensen, P. *Molec. gen. Genet.* **158**, 101–109 (1977).
8. Iida, S., Meyer, J. & Arber, W. *Plasmid* **1**, 357–365 (1978).
9. Malamy, M. H., Fiandt, M. & Szybalski, W. *Molec. gen. Genet.* **119**, 207–222 (1972).
10. Ohtsubo, H. & Ohtsubo, E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 615–619 (1978).
11. Kamp, D. *et al. Cold Spring Harb. Symp. quant. Biol.* **43**, 1159–1167 (1979).
12. Blattner, F. *et al. Science* **196**, 161–169 (1977).
13. Fiandt, M., Szybalski, W. & Malamy, M. H. *Molec. gen. Genet.* **119**, 223–231 (1972).
14. Defez, R. & De Felice, M. *Genetics* **97**, 11–25 (1981).
15. Zieg, J., Hilmen, M. & Simon, M. *Cell* **15**, 237–244 (1978).
16. Hicks, J. B. & Herskowitz, I. *Genetics* **85**, 373–393 (1977).
17. Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).
18. Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475–480 (1981).
19. Conn, E. E. *Biochem. Soc. Symp.* **38**, 277–302 (1973).
20. Schaefler, S. *J. Bact.* **93**, 254–263 (1967).
21. Szybalski, E. H. & Szybalski, W. *Gene* **7**, 217–270 (1979).
22. Studier, F. W., Rosenberg, A. H., Simon, M. N. & Dunn, J. J. *J. molec. Biol.* **135**, 917–937 (1979).
23. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
24. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).