TURN-ON OF INACTIVE GENES BY PROMOTER RECRUITMENT IN *ESCHERICHIA COLI*: INVERTED REPEATS RESULTING IN ARTIFICIAL DIVERGENT OPERONS

DANIEL CHARLIER,* YVONNE SEVERNE,* MUHAMAD ZAFARULLAH* AND NICOLAS GLANSDORFF*,†

**Microbiology, Vrije Universiteit Brussel and †Research Institute of the CERIA, B-1070 Brussels, Belgium*

Manuscript received July 25, 1982
Revised copy accepted July 15, 1983

ABSTRACT

We have characterized two rearrangements consisting of inverted repeats of the *argE* gene. The promoters (*p*) of *argE* and of *argCBH* face each other over an internal operator. The rearrangements were obained as reactivations of *argE* in a strain harboring an *argEp* deletion on a λ*darg* prophage. In both cases the repeat included *argE* and *argCBHp* on either side of a unique sequence; the result is a divergent operon in which each copy of *argCBHp* reads into the adjacent *argE* repeat. In one case, the pair of repeats adjoins the silent parental gene, forming a triplication (← → ←). The other rearrangement consists of a single *argE* palindrome, but the whole prophage is rearranged into an inverted repeat, analogous to certain λdv's. Both structures could be explained by breakage of a replication fork passing *argE* and by inaccurate rejoining of strands. The λdv-like rearrangement would result from breakage at both replication forks of a phage or prophage replicating during transient release of immunity. The triplication would imply breaking of a chromosomal replication fork, formation of a cyclic intermediate by recombination between the daughter duplex molecules and reinsertion into the parental *argE* gene. Formation of a triplication by replication errors involving appropriate strand switchings and branch migrations can not be excluded however.

THIS work is part of a study focused on DNA rearrangements able to reactivate genes deprived of their promoters. Investigations of this kind explore the cell's capacity to create or transpose gene control units; as such they probe dynamic aspects of gene evolution and may be relevant to control mechanisms involved in differentiation (NEVERS and SAEDLER 1977).

Essentially, three types of gene reactivation were known until recently: (1) insertion of a foreign sequence, *e.g.*, an IS element, in front of a gene was found in some cases to turn it on (SAEDLER *et al.* 1974; review by KLECKNER 1981) either because a promoter was created by the proper juxtaposition of sequences (BESEMER, GÖRTZ and CHARLIER 1980; GLANSDOREF, CHARLIER and ZAFARULLAH 1981; HINTON and MUSSO 1982; JAURIN and NORMARK 1983) or because the element itself carried a promoter (CHARLIER, PIETTE and GLANSDORFF 1982); (2) a deletion (MILLER *et al.* 1979) or a translocation (mutant D2 in BEEFTINK, CUNIN and GLANSDORFF 1974) fused the gene to a sequence

conducive to efficient transcription; (3) an unequal crossing over resulted in a tandem duplication placing the gene in 3′ position with respect to a foreign promoter or creating a promoter at the novel joint of the rearrangement (review by ANDERSON and ROTH (1977).

In a recent analysis of rearrangements affecting the *argECBH* divergent operon (CHARLIER *et al.* 1979) we reported on a novel, fourth type of promoter "recruitment." *ArgE* is normally transcribed from a promoter (*argEp*) facing the *argCBH* promoter over an internal operator region (ELSEVIERS *et al.* 1972; BOYEN *et al.* 1978; PIETTE *et al.* 1981, 1982; Figure 1). In a strain (Δ*sup102*) deleted for *argB*, *argC* and *argEp*, a segment comprising *argE* and the neighboring *argCBH* promoter of opposite polarity became duplicated under the form of a divergent operon where each *argE* copy appeared to be expressed from the *argCBH* promoter present on the adjacent but inverted repeat (see Figure 1, mutant AA1). The inverted repeats were found to abut the silent *argE* gene originally present in the *argEp argC argB* parental deletion strain; the resulting structure was an unstable triplication—a double palindrome—susceptible to lose of the rearrangement by looping out recombination between the original gene and the foremost copy (Figure 1). Two mechanisms were suggested (CHARLIER *et al.* 1979) to explain the formation of these structures: either DNA replication errors (Figure 2A) or formation of circles by illegitimate recombination at and behind a replication fork passing *argE*, followed by reinsertion of these circles into the chromosome at the site of the original gene (Figure 2B).

The description of this rearrangement had remained preliminary; indeed, it was uncertain whether the novel joint constituted a perfect palindrome or contained a unique sequence and whether the two inverted, face-to-face, promoters were active to the same extent. Both questions are dealt with in this paper. The first one bears on the possible mode of formation of this structure, and the second one is of interest because it concerns the mechanism of divergent transcription.

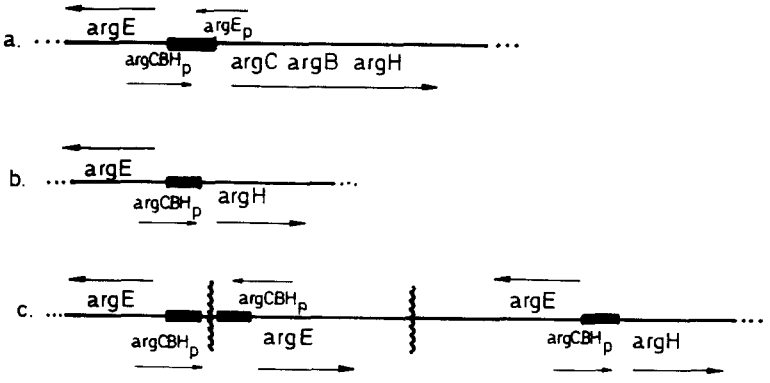

FIGURE 1.—The *argECBH* divergent operon (a), the *sup102* deletion mutant (b) and the AA1 unstable triplication (c); serrated vertical lines = novel joint; ▬ = operator region; ■ = truncated operator region in Δ*sup102* (see Figure 4 for nucleotide sequence).

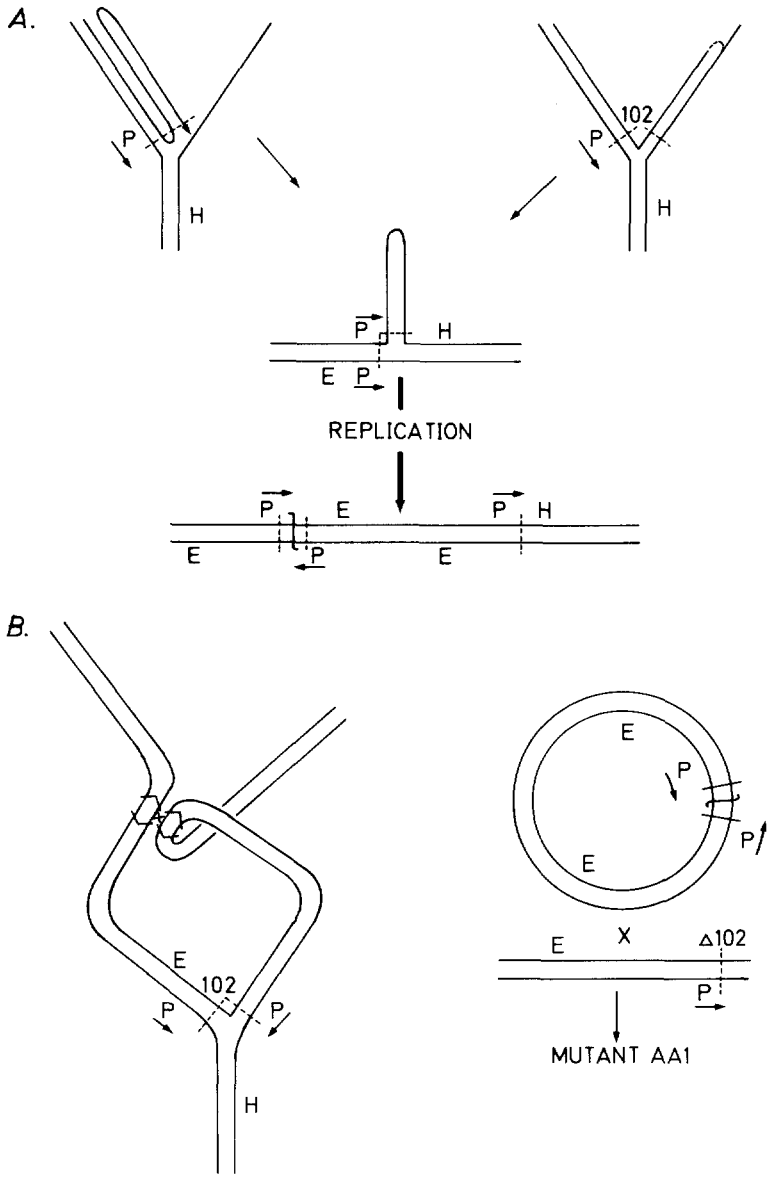

FIGURE 2.—A, Formation of a double palindrome by back and forth wandering of DNA polymerase during replication (left) or by formation of a hairpin resulting from a ligation between a newly synthesized fragment and the parental complementary region (right). B, Formation of a circular inverted repeat (left) by a break at the replication fork and a recombination event behind it; the latter would be facilitated by the presence of sequences consisting in inverted repeats. Addition to the chromosome (right) would occur by homologous recombination between two *argE* genes. P = promoter for *argCBH*; ..... = site of the *sup102* deletion; S = novel joint.

Along with these investigations on the unstable, palindromic triplications, we report on the selection of a strain (AA58) harboring stable inverted repeats consisting of only two divergent copies of the *argE* gene. The structure of the novel joint of this new type of rearrangement and its role in gene reactivation

TABLE 1

Bacterial, phage and plasmid strains

	Genotype	Source or reference
Bacteria		
P4Xλ⁻	Hfr, P4X *metB* cured of λ	BACHMAN, LOW and TAYLOR (1976)
P4XB2λ⁻	*idem*, *argR*⁻ cured of λ	This laboratory.
P4XMN42λ⁻	Hfr, P4X *metB* Δppc *argECBH* no. 42	ELSEVIERS *et al.* (1972)
P4XB2MN42λ⁻	*idem*, *argR*⁻	*idem*
P4XMN42λ⁻BF23ʳ	*idem*, phage BF23 resistant	MAZAITIS *et al.* (1976)
P4XSB167	Hfr, P4X *metB argC1*	GLANSDORFF (1965)
Phages		
λ199	*CI857 sus57 sus xis6 b515 b519*	R. WEISBERG
λ13	d *ppc argECBH bfe*	MAZAITIS *et al.* (1976)
λ13 *sup102*	d *ppc arg sup102 bfe*	*idem*
λimm434N⁻	*imm434 susN7 susN53*	R. THOMAS
BF23	wild type	R. FREDERICQ
Plasmids		
pBR322	$Ap^R Tc^R$	RODRIGUEZ *et al.* (1977)
pMC31	pBR322.H13-e[a] Ap^R Tet^R $argC^+$	CRABEEL *et al.* 1979

[a] H13-e is a 1.8-kb fragment encompassing the control region of the *argECBH* cluster.

have been analyzed in detail. A preliminary report of this work has been published in abstract form (CHARLIER *et al.* 1982).

MATERIALS AND METHODS

Abbreviations: Genetic symbols are as reported in the work of BACHMAN, LOW and TAYLOR (1976). bp = base pairs; kb = kilobase; HD = heteroduplex; EM = electron microscope; Δ = deletion.

Strains, phages and plasmids: See Table 1.

Cultures, growth conditions and media were described in GLANSDORFF (1965). Ampicillin was used at the concentration of 25 μg/ml, tetracycline and *N*-α-acetylarginine as indicated in the text.

Enzyme assays: See ELSEVIERS *et al.* (1972).

Phage propagation, induction of lysogens and preparation of heteroduplex molecules for electron microscopy were performed according to the method reported by MAZAITIS *et al.* (1976).

Preparation of plasmid DNA and transformation of E. coli was described in the work by CRABEEL *et al.* (1981).

Construction of hybrid plasmids: The technique was described by HERSCHFIELD *et al.* (1974).

Restriction enzymes and T4 ligase were ordered from New England Biolabs and used according to their instructions.

Southern blottings for identification of DNA hybrid molecules were performed as described previously (SOUTHERN 1975).

Nick translation of plasmid DNA was according to RIGBY *et al.* (1977).

DNA fragments were isolated from agarose gels as described by TABAK and FLAVELL (1978).

RESULTS

Isolation of inverted repeats

Strains AA1 (the unstable triplication mutant, CHARLIER *et al.* (1979) and AA58 (carrying a single pair of inverted repeats, see later in paper) were

isolated by the same strategy. To analyze directly the rearrangements by EM the mutants were selected from a strain (P4XMN42 Rec$^+$) carrying a deletion of the *arg* cluster and lysogenic for λ199 (a thermoinducible helper phage) and a transducing phage (λ13 *sup102*) carrying the *sup102* deletion. As the latter mutation destroys *argB*, *argC* and *argEp* without hampering *argH* (see Figure 1), the lysogen grows very slowly on *N*-α-L-acetylornithine (the natural substrate of the *argE* enzyme, see figure 3) and not at all on *N*-α-L-acetylarginine, a low affinity substrate; using acetylarginine bypasses the whole pathway, allowing the recovery of rearrangements affecting *argH*.

Independent cultures of the lysogen were grown overnight in minimal medium supplemented with 100 μg of L-arginine/ml, washed and incubated for 2 days at 32° on plates containing 100 μg/ml of acetylarginine. Colonies appeared at a frequency of about 10^{-6} the number of cells plated. As reported previously (CHARLIER *et al.* 1979; ZAFARULLAH, CHARLIER and GLANSDORFF 1981) this selection gives various types of mutants, insertions and tandem duplications usually accounting for the majority of them. Mutant AA58 appeared among 60 independent derivatives. In contrast with AA1, which came from an earlier screening (ibidem), AA58 proved unable to grow on ornithine or acetylornithine, suggesting that *argH* was impaired. In addition, although AA1 gave rise to a few percent of segregants unable to grow on acetylarginine after overnight growth in nonselective medium, AA58 proved to remain stable in those conditions.

Both AA1 and AA58, when induced, gave rise to lysates transducing the mutant phenotype at the expected high frequency.

Enzymatic characterization

A phage lysate was prepared from mutant AA58 by thermoinduction and used to infect at low multiplicity (<0.01) strain P4XMN42 (Δ*ppc argECBH*) and an *argR* (genetically derepressed) derivative, both already lysogenic for the helper phage λ199. The *ppc* gene, adjacent to *argECBH*, codes for phosphoenolpyruvate carboxylase; *ppc* strains are unable to use glucose as a sole carbon and energy source. Ppc$^+$ recombinants were selected for on minimal-glucose-arginine plates and checked for their capacity to grow on acetylarginine. Several recombinants were used to assay acetylornithinase (enzyme E) and argi-

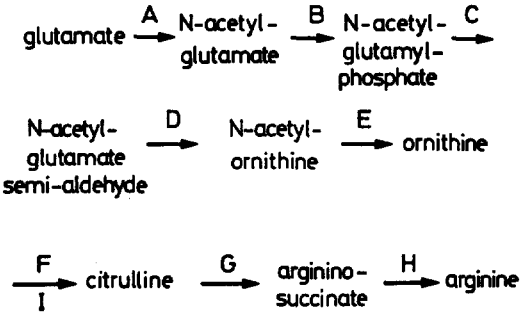

FIGURE 3.—Gene-enzyme relationship in the arginine pathway.

ninosuccinate lyase (enzyme H). The results for one of these seemingly identical recombinants are given in Table 2. Enzyme-specific activities of mutant AA1 (CHARLIER *et al.* 1979) are included for comparison.

The AA1 and AA58 mutants were found to reactivate *argE* strongly and to a similar extent. The synthesis of acetylornithinase remained five to six times repressible by arginine, suggesting that, in both mutants, the remnants of the operator region left over by the *sup102* deletion (Figures 1 and 4) were included in the rearrangement responsible for the reactivation. This will be confirmed in the "cloning" section where the DNA segments active in *argE* transcription in mutant AA1 and AA58 are analyzed in detail.

argH expression remained normal in mutant AA1, whereas in mutant AA58 no argininosuccinase activity could be detected, in keeping with the negative phenotype of this strain on media supplemented with acetylornithine or ornithine. The section on "physical analysis" will show that *argH* is actually deleted in this strain.

Physical analysis

Structure of mutant AA1: Heteroduplex mapping showed mutant AA1 to harbor overlapping *argE* palindromes: two genes keeping the original orientation flank a third, inverted exemplar [(← → ←) CHARLIER *et al.* 1979]. One aim of this work was to investigate in more details the structure of the novel joint

TABLE 2

Acetylornithinase and arginosuccinase activities

Mutant	*argR* genotype	*argE*[a]	*argH*[a]
P4X (wild type)	+	6.0	0.16
P4XB2	−	86	6.6
P4X*sup102*	+	≤0.2	1.0
P4XB2*sup102*	−	≤0.2	5.1
P4XAA1	+	67	1.1
P4XB2AA1	−	358	4.5
MN42/λ199λ13AA58	+	45	≤0.04
P4XB2Δ42/λ199λ13AA58	−	148	≤0.04

All cultures were grown in minimal medium supplemented with 100 μg of arginine/ml.

[a] Specific activities of acetylornithinase (*argE*) and argininosuccinase (*argH*) as units (micromoles of product formed per hour) per milligram of protein.

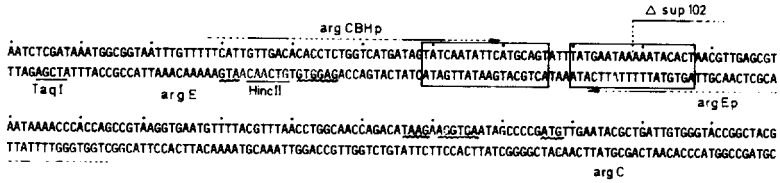

FIGURE 4.—Nucleotide sequence of the wild-type *argECBH* control region. Operators are boxed; promoters are underlined by a straight line and ribosome binding sites by a wavy line. The left hand end of the *sup102* deletion is indicated. Data from PIETTE *et al.* (1982).

between the first and the second copy, where the sequences responsible for the reactivation of *argE* are localized.

When looking at the figures obtained after reannealing λ*dargAA1* denatured DNA, we observed a single-stranded loop in some but not all molecules at the top of the 3.3-kb stem formed by intrastrand "snap-back" hybridization (Figure 5). These loops were only a few hundred nucleotides long so that they presented a variable appearance and were sometimes not discernible because of collapsing; this explains our failure to observe them in a previous study (CHARLIER *et al.* 1979). Whenever visible, this loop was present only at the end of the stem made between the first and the second copy, that is at 4.9 kb from the end of the autoannealed single-stranded molecules or, in heteroduplexes with λ13, in the stem formed to the left of the loop marking the *sup102* deletion. Never could a clear single-stranded loop be observed at the other stem position. Therefore, a unique sequence of a few hundred base pairs was present at one of the novel joints but not at the other one.

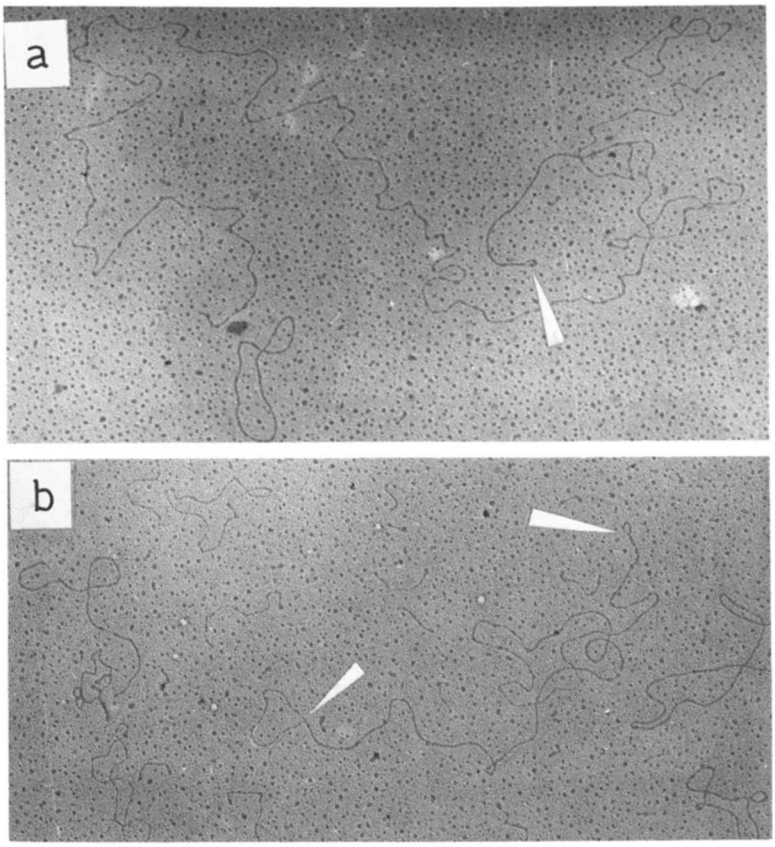

FIGURE 5.—Electron micrograph of autoannealed single strands of (a) λ13AA1 showing a small single-stranded loop at the top of the double-stranded 3.2-kb long snap back structure. (b) λ13AA58 shows a 19-kb long double-stranded segment with a small loop of a few hundred base pairs at one end and a large one of 2.7 kb at the other end.

In λ13/λAA1 heteroduplexes (CHARLIER *et al.* 1979, Figures 2 and 7), the base of the stem formed by the second and third copies was found some 300–400 bp to the right of the loop marking the *sup102* deletion. In this short region, no single-stranded DNA could be detected. Therefore, the unique sequence now found at the novel joint must contain DNA normally present to the right of the deletion, *i.e.*, part of *argH*. This was confirmed by restriction analysis and found to apply to mutant AA58 as well (see *Cloning and comparison of the novel joints in mutant AA1 and AA58*).

Structure of mutant AA58: Lysates of mutant AA58 gave two visible bands when centrifuged to equilibrium on a CsCl density gradient. The lower one contained plaque-forming helper phage, the upper one phage conferring the AA58 phenotype upon transduction.

Denatured and reannealed molecules of λ13AA58 showed a long double-stranded snap-back structure of 19 kb with a 2.7-kb long single-stranded loop on one side and another loop, a few hundred base pairs long and of variable appearance, on the other side (see Figure 5). The DNA extracted from the capsids was a 41-kb long linear duplex so that the observed structures must be due to snap-back hybridization of single strands, the ends of the molecule being so close to each other that the junction could not be discerned by EM. To map the endpoints, autoannealed molecules were digested with the single strand-specific S1 nuclease and the products analyzed by EM and agarose gel electrophoresis. Three distinct classes of molecules were observed in the microscope; their lengths were 20.0, 10.5 and 8.5 kb, respectively, in good agreement with the electrophoretic pattern which gave bands of 19.5, 10.2 and 8.6 kb. The largest fragment should correspond to the stem (see Figure 5b) cleared of the two single-stranded loops but not yet attacked at the internal site where the two ends of the molecule come in close apposition.

The respective position of the two inverted repeats with respect to the single-stranded loop could be deduced from the observation of autoannealed broken molecules occurring in the preparation. Broken molecules which lacked a segment at only one end showed a single-stranded region of variable length, leaving only one of the two inverted repeats intact.

In summary, the data showed λ13AA58 to consist of two extensive inverted repeats (10.2 and 8.6 kb long) with unique sequences in the middle (2.7kb and a few hundred base pairs long, respectively). Most of the genome of the transducing phage was, therefore, duplicated; consequently, much of the parental DNA must have been deleted since, in spite of these large duplications, the phage was still packageable, with a duplex length of 41 kb.

Restriction endonuclease mapping: Heteroduplex mapping of λ13AA58 was made difficult by rapid intrastrand hybridization of the long inverted repeats. To localize the different bacterial and λ genes present on λ13AA58, the electrophoretic pattern of restriction fragments of the phage was compared with these of λ*sup102* and λ199 (CRABEEL *et al.* 1977; MORAN *et al.* 1977). The enzymes chosen for single and double digestions were *Hind*III, *Bam*HI and *Eco*RI, since they could yield only a limited number of fragments. Combining

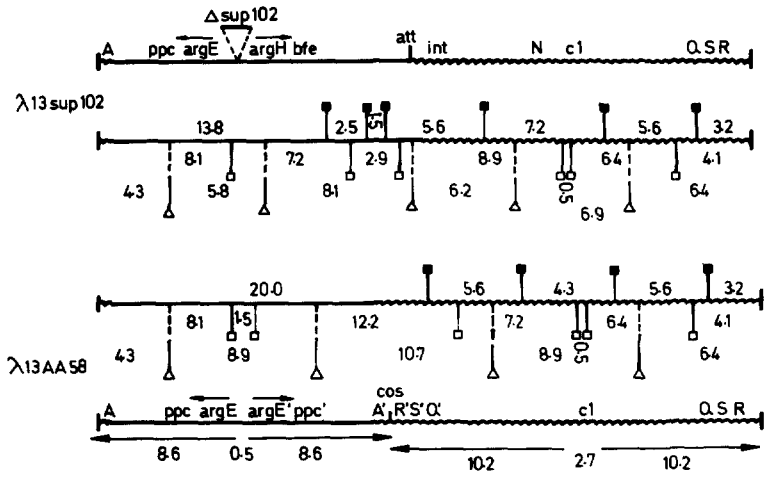

FIGURE 6.—Restriction endonuclease and vegetative maps of λ13*sup102* and λ13AA58. Lengths are in kilobases. ——□ = *Hind*III restriction sites; ——■ = *Eco*RI sites; ——◁ = *Bam*HI restriction sites. Inverted repeats are indicated by arrows below the map of λ13AA58.

the results of these investigations and those of the previous EM analysis gave the physical map presented on Figure 6.

The map implies that λ13AA58 has no *att* site but possesses two *cos* sites. Phages with two *cos* sites have already been described by EMMONS (1974); he observed that the packaging mechanism gave rise to phage DNA molecules of different lengths depending on which *cos* site the *ter* nuclease acted. In the case of λ13AA58, where the whole genome has been duplicated, the situation around the two *cos* sites is identical, and no heterogeneity of packaging products is expected.

The map also implies that the *argH* and *bfe* genes are absent in λ13AA58, λ13AA58 is a $N^-C_I^+$ phage and the 1.5-kb *Hind*III fragment is delimited by two *Hind*III sites localized in inverted *argE* repeats flanking the novel joint.

λ13AA58 lacks argH *and* bfe*:* The absence of a functional *argH* gene was already attested by the enzyme assay reported in Table 2 and the growth test on different media. The *bfe* gene confers sensitivity toward bacteriophage BF23; the absence of that gene should result in resistance to the phage, a recessive trait. A Bfe^r derivative of MN42λ^- was lysogenized with λ13AA58 and λ199. Cultures of this strain, of MN42λ^- (negative control) and of P4X (positive control), were infected in the early exponential phase with phage BF23. No lysis could be observed with the λ13AA58 lysogen, whereas a positive answer was obtained with P4X.

λ13AA58 lacks the N *gene*: To determine whether the *N*λ gene was lacking in λ13AA58, a genetic complementation test was performed between λ13AA58 and the heteroimmune phage λ*imm434 susN7 susN53*. No phage production could be observed when a λ13AA58 lysogen was superinfected with the double N^- mutant phage whereas a control experiment performed with a λ13sup102 (N^+) lysogen gave the opposite result.

A $N^-C_I^-cro^+$ phage can establish itself as a relatively stable plasmid in a λ^- host (SIGNER 1969). However, a λ $N^-C_I^+$ cro^+ phage can produce enough repressor to repress plasmid propagation (KLECKNER and SIGNER 1977); this should be the case with λ13AA58. Indeed, infection with purified λ13AA58 of λ^+ or λ^-, respectively, Rec^+ or *recA* derivatives of a strain deleted for the *arg* region showed that the only way this phage could stably establish itself was recombinational insertion into the chromosome of a Rec^+ $(\lambda 199)^+$ host.

The 1.5-kb Hind*III fragment carries the novel joint between the two divergent* argE *repeats:* The structure drawn in Figure 6 predicted the two *Hind*III sites bordering the 1.5-kb fragment to be localized in divergent *argE* repeats flanking a region comprising the novel joint and the unique sequence, including the sites controlling acetylornithinase synthesis. This prediction was borne out by SOUTHERN (1975) blotting experiments. Plasmid PMC31, a pBR322 derivative containing a 1.8-kb *Hind*III fragment spanning the wild-type *argECBH* control region (CRABEEL *et al.* 1979), was labeled by nick translation and used as a probe to hybridize with a *Hind*III digest of λ13AA58. The only band revealed in the autoradiogram was indeed the 1.5-kb fragment. This result did not yet prove that a complete *argCBH* promoter was present on both sides of the novel joint, but this was established by the experiments described in the next section.

Cloning and comparison of the novel joints in mutant AA1 and AA58

In the next step we determined the fine structure of the novel joints in both mutants and investigated whether the two promoters facing each other over the unique sequence were equally active in transcription.

A 1.5-kb fragment flanked by *Hind*III sites located in divergent *argE* repeats could be recovered from λ13AA1 as well as from λ13AA58 (described before). From both phages, the fragment was isolated and cloned into the unique *Hind*III site of plasmid pBR322; the recombinant plasmid was recovered and maintained by selecting for resistance toward ampicillin.

Cloning into the *Hind*III site of pBR322 should destroy the promoter for the gene confering resistance toward tetracycline (RODRIGUEZ *et al.* 1977). This would result in tet^s clones unless the fragment inserted contained a promoter geared in the appropriate orientation. Amp^r clones were screened on concentrations of tetracycline ranging from 2.5–30 μg/ml. Only two classes could be distinguished whether the fragment came from AA1 or AA58; they were resistant up to 5 and 15 μg/ml, respectively. Several representatives of the two kinds of clones were shown to contain the 1.5-kb *Hind*III fragment by restriction enzyme analysis.

Plasmids pCD1-117 (from λ13AA1) and pCD58-26 (from λ13AA58) conferred resistance to 5 μg; pCD1-163 (from λ13AA1) and pCD58-101 (from λ13AA58) were resistant to 15 μg/ml.

As two different levels of resistance were obtained with the same fragment, one could expect this difference to be due to the orientation of the inserted fragment. *Hind*III restriction of pCD1-117 or pCD58-26 followed by religation of the mixture of fragments thus obtained gave rise in each case to the same two classes of resistant strains after transformation. This already indicated that the expression was indeed orientation dependent.

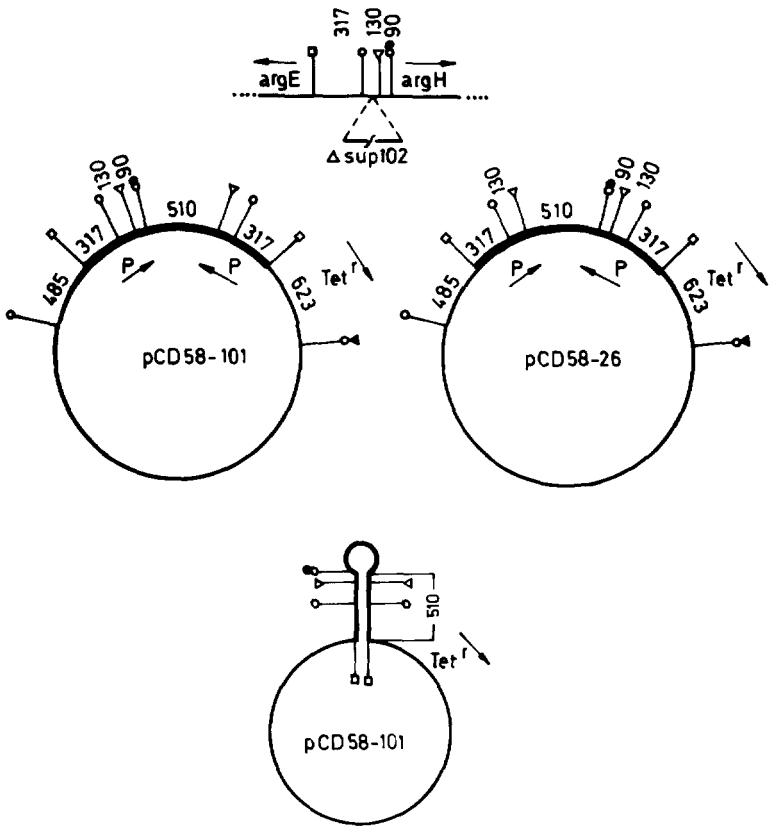

FIGURE 7.—Restriction maps of the arginine region in the vicinity of the *sup102* deletion (top line) and of plasmids pCD58-101 and pCD58-26 bearing the same 1.5-kb *Hind*III fragment cloned from λ13AA58 but in opposite orientation (thick line). pCD1-163 (from λ13AA1, not shown) is similar to pCD58-101 except that the most rightward of the two *Hinf*I sites is missing (see text); pCD1-117 (also from λ13AA1) is similar to pCD58-26 except for the most leftward *Hinf*I site. Lengths are given in base pairs. The *Hinc*II site at 317 bp to the right of the *Hind*III site is in the promoter for *argCBH* (Figure 3; Piette *et al.* 1982). ——□ = *Hind*III restriction sites; ——○ = *Hinc*II sites; ——◁ = *Hinf*I sites; ——● = *Hpa*I sites; ——◀ = *Sal*I sites. The *Hinf*I sites are only shown on the bacterial insert and not on the pBR322 sequences. The lower drawing illustrates an HD molecule obtained by reannealing single-stranded DNA from plasmid pCD58-101, as described in the text.

To determine this orientation and to analyze in more detail the sequences present in the cloned fragment of different plasmids, *Hinc*II and *Hpa*I restriction maps were constructed and compared with the corresponding map of the arginine region in the vicinity of the *sup102* deletion (Figure 7; PIETTE *et al.* 1982). To the right of the *Hind*III site in *argE*, 317 bp, there is a *Hinc*II cut localized in the promoter for *argCBH*; 217 bp further—and now to the right of *sup102* at 170 bp—there is a second *Hinc*II site which is also recognized by *Hpa*I; finally, there is a *Hinf*I cut 444 bp to the right of the *Hind*III cut, thus between the two *Hinc*II sites and 75 bp beyond *sup102*. *Hinc*II restriction of the four plasmids yielded five fragments. The relevant sites were localized by double *Hind*III-*Hinc*II, *Hind*III-*Hpa*I and *Sal*I-*Hpa*I digests (see Figure 7).

Two *Hinc*II cuts proved to occupy symmetrical positions in the inverted repeats at 317 bp from the *Hind*III sites bordering the cloned fragment. The position of the *Hpa*I/*Hinc*II cut was found to be asymmetrical, which allowed us to determine the orientation of the 1.5-kb *Hind*III fragment cloned in the plasmid.

In pCD1-163 and pCD58-101, the *Hpa*I cut was located on the far side of the cloned fragment with respect to the *tet* gene, whereas in the two other plasmids the same site was closer to the *tet* gene (Figure 7).

To determine whether two complete *argCBH* promoters were present on both sides of the unique sequence the purified 1.5-kb *Hind*III fragment from mutants AA1 and AA58 were restricted by the *Hinf*I endonuclease, which cuts in *argH*, well beyond the *sup102* deletion (Figure 7), The *Hinf*I pattern of mutant AA58 gave two fragments estimated at 610 ± 20 bp and 440 ± 20 bp, respectively; the latter band was of a greater intensity, due to enzymatic cutting at two symmetrically located sites 444 bp from either end of the 1.5-kb fragment. Two complete promoters are, therefore, present on that piece of DNA.

This conclusion was confirmed by the experiment outlined in Figure 7: DNA from plasmid pCD58-101 was denatured and allowed to reanneal for 10 min at 37° to allow the inverted repeats flanking the unique sequence to form duplex molecules; S1 nuclease was then used to digest single-stranded DNA, and the duplex fragment was submitted to electrophoresis on a 5% polyacrylamide gel in parallel to fragments of standard length in order to estimate its size. The fragment was found to be 510 ± 10 bp long; therefore, the inverted repeat terminates some 30 bp left of the *Hpa*I site, thus well beyond the control region.

The *Hinf*I restriction pattern of AA1 showed two bands estimated at 1100 and 440 bp, respectively, thus indicating the presence of only one *Hinf*I site. In restriction analysis of the 1.5-kb *Hind*III fragment from both AA58 and AA1, the internal *Hinc*II fragments migrated so similarly that the difference between them should not exceed 50 bp; therefore in AA1, also, the two promoters might be complete as well. This was ascertained according to the scheme outlined in the previous paragraph and in Figure 7 with DNA from plasmid pCD1-117. After denaturation, reannealing and S1 nuclease treatment, a fragment of 440 bp was obtained; therefore, in mutant AA1, the inverted repeat falls about 10 bp short of the *Hinf*I site, still beyond *sup102*, and the two facing promoters are intact.

From these measurements of the inverted repeats flanking the unique sequence, the length of the latter could be estimated to be 480 bp for AA58 (1.5–2 × 0.510 kb) and 620 (1.5–2 × 0.440 kb) for AA1. A detailed map of the two rearrangements can be found in Figure 8.

It was already known (see *Structure of mutant AA1*) that the unique sequence present in this rearrangement was part of *argH*. By HD analysis we could show that the unique sequences of both AA1 and AA58 are homologous. DNA from plasmids pCD58-101 and pCD1-117 were denatured and reannealed together. Since the two plasmids carry the 1.5-kb *Hind*III fragment in opposite

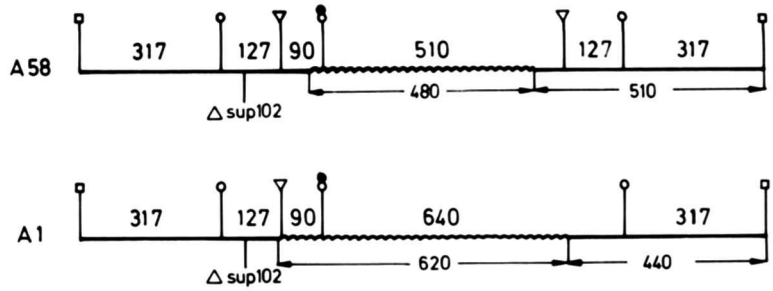

FIGURE 8.—Structure of the novel joints of mutant AA1 and AA58. Serrated line = unique sequence; —□ = *Hind*III restriction sites; —○ = *Hinc*II restriction sites; —◁ = *Hinf*I restriction sites; —● = *Hpa*I sites; —◀ = *Sal*I sites.

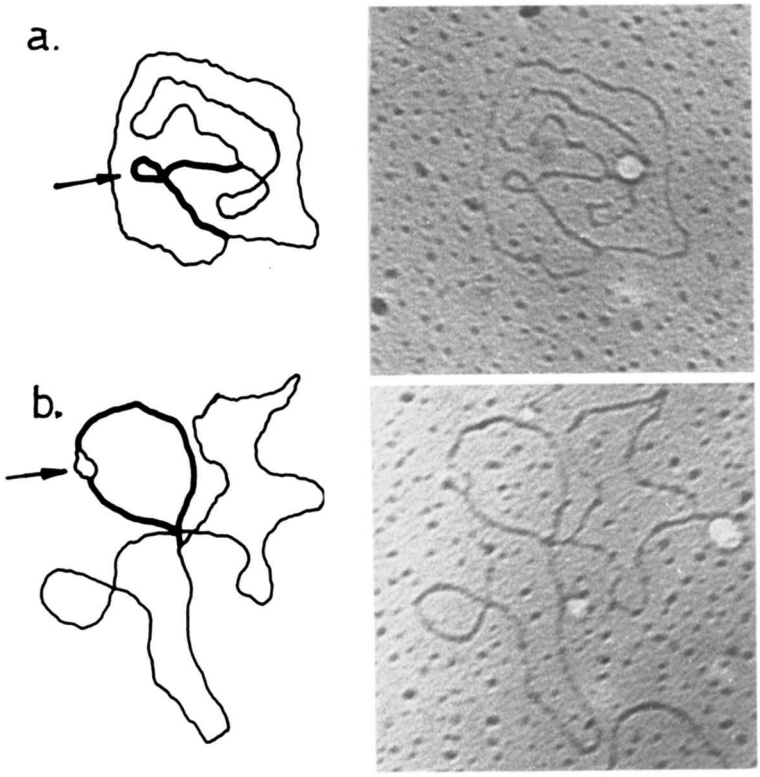

FIGURE 9.—HD structures obtained after reannealing denatured DNA molecules from plasmids pCD58-101 and pCD1-117 which, respectively, carry the 1.5-kb *Hind*III fragment from mutants AA58 and AA1, and this in opposite orientations (see text for further comments).

orientations we could expect either of the following structures if the unique sequences were homologous: two single-stranded plasmid DNA molecules hybridizing at the level of their unique sequences only (Figure 9a, arrow), the inverted repeats of each individual molecule having snapped back on themselves to form double-stranded stems; two single-stranded molecules hybridizing over most of a 1.5-kb long segment with a short "eye" (Figure 9b, arrow),

which we expect to be formed in the region of the *Hinf*I site, where the two strands lack homology for about 100 bp (see Figure 8 for an expanded map of the two rearrangements).

In the plasmids conferring the higher degree of resistance toward tetracycline (pCD1-163 and pCD58-101) the unique sequence is oriented with respect to the *tet* gene in the way it is normally read in the *arg* cluster. RNA polymerase molecules initiating at the other promoter would transcribe the antisense strand of the unique sequence. This difference is probably the cause for the difference in strength of the two otherwise identical promoters present on the 1.5-kb fragment (see DISCUSSION).

DISCUSSION

We have investigated how a silent gene can be turned-on by a nearby promoter in opposite polarity. The process involved consists of *de novo* formation of a divergent operon where two facing copies of the "recruited" promoter read into adjacent but inverted repeats of the reactivated gene.

The duplication of genes in direct repeats was until recently the only type of tandem duplication known to geneticists. Their formation is usually explained by unequal crossovers, whether they be legitimate (*i.e.*, depending on recombination between homologous segments) or illegitimate (reviewed by ANDERSON and ROTH 1977). Obviously, the same scheme could not apply to explain the formation of inverted repeats. We will successively consider mutants AA1 (the unstable double palindrome ← → ←) and AA58 (the stable single pair of repeats ← →).

Figure 2 recalls the two models we had in mind (CHARLIER *et al.* 1979) to explain mutant AA1 when it was first described: either errors in DNA replication involving template switching (Figure 2A) or the recombinational model (Figure 2B); the latter would involve a break in the replication fork, another recombination upstream (*e.g.*, between inverted repeats) and reinsertion of the so-formed circle in the chromosome by homologous recombination within *argE* DNA. The latter model postulates a kind of event (break at the replication fork with respective joining of parental and progeny strands which has been assumed to produce the symmetric type of λdv (CHOW, DAVIDSON and BERG 1974): in the case of λdv, however, a θ replication form would have to break at both replication forks. As pictured on Figure 10, this type of illegitimate recombination at a replication fork would involve the inclusion of a unique sequence at the novel joint if the two parental strands were out of register at the time the break occurred. This unique sequence could come from either one of the strands; their presence has been observed in several symmetric λdv's and has now been detected in mutant AA1 as well.

This similarity between the novel joint of some λdv's and of mutant AA1 is presently the strongest argument in favor of the recombinational model. The demonstration that such rearrangements do not take place in *recA* strains would certainly strengthen the case since homologous recombination within *argE* is required as the last step. Fortunately, a more positive approach to the problem is also possible and is presently being pursued: were this model correct it should

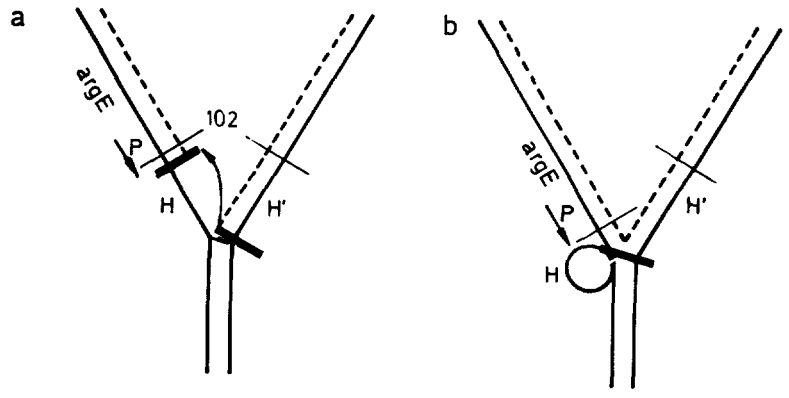

FIGURE 10.—Formation of an inverted repeat of *argE* with inclusion of a unique sequence (H and H′ for part of *argH*) at the novel joint by illegitimate recombination at a replication fork; a, illegitimate recombination occurs before abnormal replication; b, abnormal replication occurs before illegitimate recombination. Symbols:102 = deletion *sup102*; ■ = break; arrow = joining of strands; P = promoter for *argCBH*.

be possible to "trap" the postulated cyclic intermediate by homologous recombination with *argE* DNA inserted in a plasmid or an episome; the latter could then be recovered after screening for those *argE* reactivations which could be transferred by transformation or conjugation. Such experiments will be useful because, at present, schemes involving replication errors cannot be excluded. Simple template switching would certainly not account for the presence of the unique sequence at the novel joint; rather, a series of template switchings and branch migrations would be required; several alternative pathways are depicted on Figure 11. It is difficult to estimate how likely these sequences of events would be; template switching occurs *in vitro* (see KORNBERG 1980) and probably *in vivo* (see discussion by RIPLEY 1982 and BOLIVAR *et al.* (1977) but it may be questioned whether the 4–600-nucleotide long and looped-out unique sequence postulated by the model in alternative a and c (Figure 11) (the latter one inspired from RIPLEY 1982) would be stable enough in order to last until completion of the rearrangement. Also, the very length (3.0 kb) of the hairpin that should have been formed by template switching and possibly displaced by branch migration (as on Figure 11 c) is considerable. An alternative that does not require loop stability and involves less switching and migration than the former ones is depicted on Figure 11b; it involves branch migration over the length of the unique sequence, template switching and formation of a hairpin resulting from ligation between the newly synthesized fragment and the parental complementary region; the next replication round would complete the rearrangement.

As we have already remarked (CHARLIER *et al.* 1979) the occurrence of palindromes would increase the probability of the back and forth template switching depicted in Figure 11a; however, as noted, palindromes would also serve the recombinational model by allowing the crossing over required upstream of the replication fork to form the cyclic intermediate (Figure 2B). A

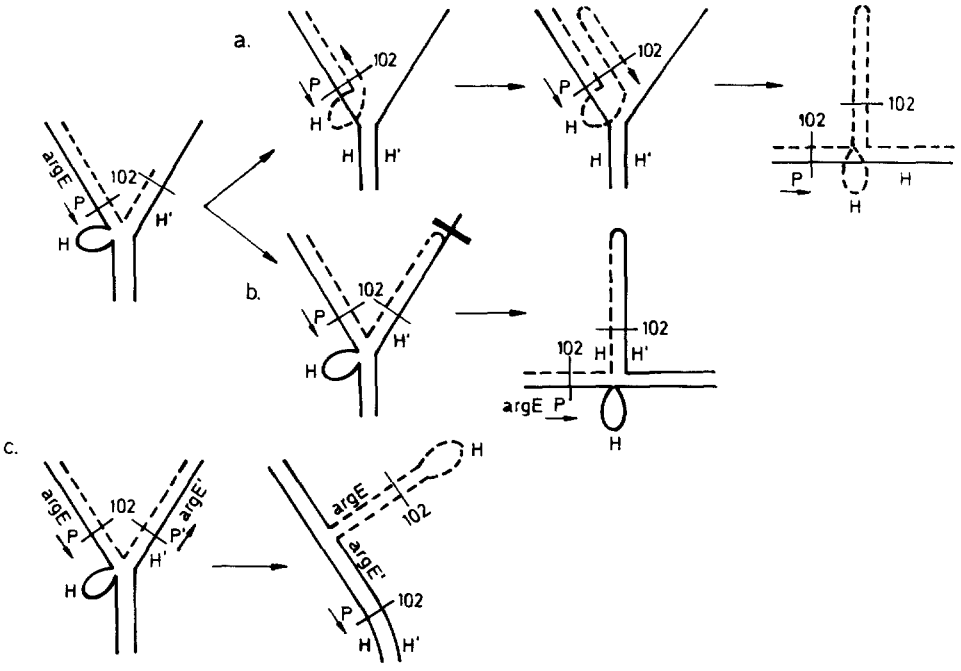

FIGURE 11.—Errors in DNA replication leading to the formation of a palindromic triplication with inclusion of a unique sequence (H, H', for part of *argH*) at the novel joint between two inverted *argE* repeats. Symbols as in Figure 10. a, Displacement of sequence H and template switching followed by further branch migration of H and two successive (back and forth) template switches. Rearrangements completed at next replication. b, Displacement of sequence H and template switching along the complementary parental strand followed by hairpin closure at the site indicated. Rearrangements completed at next replication. c, Displacement of sequence H and template switching along the complementary parental strand followed by branch migration of *argE*. Rearrangements completed at next replication.

good criterion to check the validity of an explanation based on replication errors instead of recombination would be to determine whether such rearrangements do occur in a *recA* strain. Extensive screening will be required in view of the variety of rearrangements (duplication, insertions, deletions) found when looking for reactivation of *argE* in the *sup102* strain. Of course, both models could be correct, each one accounting for a particular class of events.

The case of mutant AA58 appears more straightforward than that of AA1, because of the great structural resemblance between the whole AA58 prophage and the palindromic λdv's. In this case, only one type of recombinational event had to be postulated: a break occurring at both forks of a divergently replicating λ*darg* (Figure 12). This could have occurred if the λ*dargsup102* prophage had been excised by homologous recombination from the λ*darg*-λ prophage tandem present in the chromosome and had subsequently undergone partial replication in the cytoplasm during a transient release of immunity. The AA58 rearrangement would have been completed by reinsertion into a chromosome carrying a λ prophage, possibly the very same one from which λ*sup102* would have been excised before undergoing the rearrangement. Spontaneous

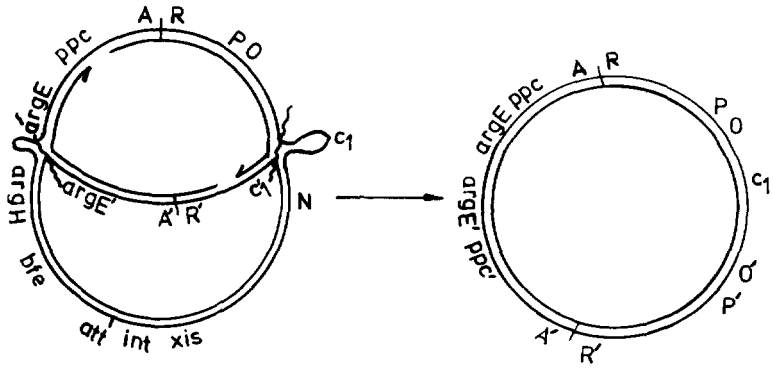

FIGURE 12.—Formation of a single, stable *argE* inverted repeat by breakage at the two replication forks of a temporarily excised λ13*sup102* phage; inaccurate rejoining of parental and progeny strands may result in two inverted repeats separate by unique sequences (H and C_1). Reinsertion of the circular intermediate into the chromosome by homologous recombination between λ genes and the helper λ199 prophage would yield mutant AA58. A nondepicted alternative would involve initiation of prophage replication on the chromosome (see text).

excision from prophage in tandem occurs frequently, as expected (C. DAMBLY, personal communication), but it is not strictly required in the present case; indeed, initiation of divergent prophage replication promoted by a transient release of immunity could happen on the chromosome itself, then be followed by illegitimate recombination at both replication forks prior to reinsertion of the cyclic intermediate into the chromosome. From our own estimates, spontaneous induction accounts for the presence of about 10^5 phage particles per 5×10^8 cells of the $(\lambda 199)^+$ (λ*dargsup102*) lysogen; the frequency of abortive inductions could be much higher.

Regarding gene expression in mutants AA1 and AA58, the main point of this paper is to have shown that there is a unique sequence at the junction of the two repeats, flanked by two complete and facing copies of the *argCBH* promoter. The Pribnow box of *argCBH* overlaps the operator region by at least four nucleotides (CUNIN *et al.* 1982;PIETTE *et al.* 1982; Figure 3); part of the operator has been removed by the *sup102* deletion, but the remains are sufficient to ensure a five- to ten-fold amplitude of variation in the expression of *argH* (instead of 60-fold in the wild type). Since *argE* is regulated in both mutants to the same extent as *argH* in the parental *sup102* strain, no other sequences than the two *argCBH* promoters appear responsible for acetylornithinase synthesis in those strains.

The unique sequence consists of about four to 600 nucleotides which are part of the *argH* gene on the basis of current DNA sequence data (PIETTE *et al.* 1982; CHARLIER, PIETTE and GLANSDORFF 1982). Thus, only one of the two facing promoters read into a stretch of "sense" DNA before reaching the adjacent *argE* repeat. This asymmetry provides a reasonable explanation for the fact that one of the promoters yields more gene product than the other one; indeed, considering that the antisense strand might contain sequences interferring either with messenger elongation, with correct RNA translation or with both, it comes as no surprise that it is the promoter reading into sense

DNA (the leftward one in pCD58-101 and pCD1-163) that determines the higher level of gene expression. If the mechanism just proposed for the inclusion of the unique sequence in the final structure is correct, either the leftward or the rightward promoter could in principle had been the one followed by sense DNA.

Rearrangements as discussed in this paper could have played a role in the genesis of divergent operons and, in general, in the formation of inverted repeats of some length; they could also be the cause of certain map symmetries (CHARLIER *et al.* 1979). In eukaryotes, where several origins of divergent replication may be present on each chromosome, many genes would become bracketed by replication forks during DNA synthesis and rearrangements like the one displayed by AA58 could be more common than in *E. coli*.

This work was supported by the Belgian Foundation for Collective and Fundamental Research (grant 2.9005.79). M. Z. acknowledges a fellowship from the Belgian-Pakistanese Cultural Agreement. We thank CHRISTINE DAMBLY and ARIANE TOUSSAINT for helpful discussion.

LITERATURE CITED

ANDERSON, R. P. AND J. R. ROTH, 1977 Tandem genetic duplications in phage and bacteria. Annu. Rev. Microbiol. **31:** 473–505.

BACHMANN, B. J., B. LOW and A. L. TAYLOR, 1976 Recalibrated linkage map of *Escherichia coli* K-12. Bacteriol. Rev. **40:** 116–170.

BEEFTINK, F., R. CUNIN and N. GLANSDORFF, 1974 Arginine gene duplications in recombination proficient and deficient strains of *Escherichia coli* K-12. Mol. Gen. Genet. **132:** 241–253.

BESEMER, J., I. GÖRTZ and D. CHARLIER, 1980 Deletions and DNA rearrangements within the transposable element IS2: a model for the creation of palindromic DNA by DNA repair synthesis. Nucleic Acids Res. **8:** 5825–5833.

BOLIVAR, F., M. C. BETLACH, H. L. HEYNEKER, J. SHINE, R. L. RODRIGUEZ and H. W. BOYER, 1977 Origin of replication of pBR345 plasmid DNA. Proc. Natl. Acad. Sci. USA **74:** 5265–5269.

BOYEN, A., D. CHARLIER, M. CRABEEL, R. CUNIN, S. PALCHAUDHURI and N. GLANSDORFF, 1978 Studies on the control regions of the bipolar *argECBH* operon of *Escherichia coli*. I. Effect of regulatory mutations and IS2 insertions. Mol. Gen. Genet. **161:** 185–196.

CHARLIER, D., M. CRABEEL, R. CUNIN and N. GLANSDORFF, 1979 Tandem and inverted repeats of arginine genes in *Escherichia coli*: structural and evolutionary considerations. Mol. Gen. Genet. **174:** 75–88.

CHARLIER, D., J. PIETTE and N. GLANSDORFF, 1982 IS3 can function as a mobile promoter in *E. coli*. Nucleic Acids Res. **10:** 5935–5948.

CHARLIER, D., M. ZAFARULLAH, Y. SEVERNE and N. GLANSDORFF, 1982 *ArgE* palindromes in *Escherichia coli*. Arch. Int. Physiol. Biochim. **90:** B6–B7.

CHOW, L., N. DAVIDSON and D. BERG, 1974 Electron microscope study of the structure of λdv DNAs. J. Mol. Biol. **86:** 69–89.

CRABEEL, M., D. CHARLIER, R. CUNIN and N. GLANSDORFF, 1979 Cloning and endonuclease restriction analysis of *argF* and of the control region of *argECBH* bipolar operon in *Escherichia coli*. Gene **5:** 207–231.

CRABEEL, M., D. CHARLIER, N. GLANSDORFF, S. PALCHAUDHURI and W.K. MAAS, 1977 Studies on the bipolar *argECBH* operon of *E. coli*: characterization of restriction endonuclease fragments obtained from λ*dargECBH* transducing phages and a colE1 *argECBH* plasmid. Mol. Gen. Genet. **151:** 161–168.

CRABEEL, M., F. MESSENGUY, F. LACROUTE and N. GLANSDORFF, 1981 Cloning *arg3*, the gene for ornithine carbamoyltransferase from *Saccharomyces cerevisiae*: expression in *Escherichia coli* requires secondary mutations; production of plasmid β-lactamase. Proc. Natl. Acad. Sci. USA **78:** 5026–5030.

CUNIN, R., A. BOYEN, J. PIETTE, M. CRABEEL and N. GLANSDORFF, 1982 Superposition of genetic sites in the regulatory region of the bipolar *argECBH* operon of *Escherichia coli*. Ann. Microbiol. (Paris) **133A:** 235–241.

ELSEVIERS, D., R. CUNIN, N. GLANSDORFF, S. BAUMBERG and E. ASCHROFT, 1972 Control regions within the *argECBH* cluster of *Escherichia coli* K-12. Mol. Gen. Genet. **117:** 349–366.

EMMONS, S. W., 1974 Bacteriophage lambda derivatives carrying two copies of the cohesive end site. J. Mol. Biol. **83:** 511–525.

GLANSDORFF, N., 1965 Topography of cotransducible arginine mutations in *Escherichia coli* K-12. Genetics **51:** 167–179.

GLANSDORFF, N., D. CHARLIER and M. ZAFARULLAH, 1981 Activation of gene expression by IS2 and IS3. Cold Spring Harbor Symp. Quant. Biol. **45:** 153–156.

HERSCHFIELD, V., H. BOYER, C. YANOFSKY, M. LOVETT and D. HELINSKI, 1974 Plasmid colE1 as a molecular vehicle for cloning and amplification of DNA. Proc. Natl. Acad. Sci. USA **71:** 3455–3459.

HINTON, D. and R. MUSSO, 1982 Transcription initiation sites within an IS2 insertion in a Gal-constitutive mutant of *Escherichia coli*. Nucleic Acids Res. **10:** 5015–5031

JAURIN, B., and S. NORMARK, 1983 Insertion of IS2 creates a novel *ampC* promoter in *Escherichia coli*. Cell **32:** 809–816.

KLECKNER, N., 1981 Transposable elements in prokaryotes. Annu. Rev. Genet. **15:** 341–404

KLECKNER, N. AND E. SIGNER, 1977 Genetic characterization of plasmid formation by N^- mutants of bacteriophage λ. Virology **79:** 160–173.

KORNBERG, A., 1980 *DNA Replication*. pp. 153–154. Freeman, San Fransisco.

MAZAITIS, A., S. PALCHAUDHURI, N. GLANSDORFF and W. K. MAAS, 1976 Isolation and characterization of λ*dargECBH* transducing phages and heteroduplex analysis of the *argECBH* cluster. Mol. Gen. Genet. **143:** 185–196.

MILLER, J. M., W. S. REZNIKOFF, A. E. SIVERSTONE, K. IPPEN, E. R. SIGNER and J. R. BECKWITH, 1979 Fusion of the lac and trp regions of the *Escherichia coli* chromosome. J. Bacteriol. **101:** 1273–1279.

MORAN, A., A. MAZAITIS, R. VOGEL and H. VOGEL, 1977 BamHI, HindIII and EcoRI restriction endonuclease cleavage sites in the *argECBH* region of the *Escherichia coli* chromosome. Pp. 273–278. In: *Nucleic Acid-Protein Recognition*, Edited by H. VOGEL, Academic Press, New York.

NEVERS, P. and H. SAEDLER, 1977 Transposable genetic elements as agents of gene instability and chromosomal rearrangements. Nature **268:** 109–115.

PIETTE, J., A. BOYEN, M. CRABEEL, R. CUNIN and N. GLANSDORFF, 1981 Characterization of promoter locus in the control region of the *argECBH* bipolar operon of *Escherichia coli*. Arch. Int. Physiol. Biochim. **89:** B126–B127.

PIETTE, J., R. CUNIN, A. BOYEN, D. CHARLIER, M. CRABEEL, F. VAN VLIET, N. GLANSDORFF, C. SQUIRES and C. L. SQUIRES, 1982 The regulatory region of the divergent *argECBH* operon in *Escherichia coli* K-12. Nucleic Acids Res. **10:** 8031–8048.

RIGBY, P., M. DIECKMANN, C. RHODES and P. BERG, 1977 Labelling deoxynucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113:** 237–251.

RIPLEY, L. S., 1982 Model for the participation of quasi-palindromic DNA sequences in frameshift mutations. Proc. Natl. Acad. Sci. USA **79:** 4128–4132.

RODRIGUEZ, R. L., R. TAIT, J. SHIVE, F. BOLIVAR, H. HEYNEKER, M. BETLACH and H. W. BOYER, 1977 Characterization of tetracycline and ampicilline resistant plasmid cloning vehicles. pp. 73–84. In: *Molecular Cloning of Recombinant DNA*, Edited by W. A. SCOTT and R. WERNER, Vol. 13, Miami Winter Symposia University of Miami.

SAEDLER, H., H. J. REIF, S. HU and N. DAVIDSON, 1974 IS2, a genetic element for turn-off and turn-on of gene activity in *E. coli*. Mol. Gen. Genet. **132:** 265–289.

SIGNER, E., 1969 Plasmid formation: a new mode of lysogeny by phage lambda. Nature **223:** 158–160.

SOUTHERN, E. M., 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 503–517.

TABAK, H. F. and R. A. FLAVELL, 1978 A method for the recovery of DNA from agarose gels. Nucleic Acids Res. **5:** 2321–2332.

ZAFARULLAH, M., D. CHARLIER and N. GLANSDORFF, 1981 Insertion of IS3 can "turn-on" a silent gene in *Escherichia coli*. J. Bacteriol. **146:** 415–417.

Corresponding editor: D. BOTSTEIN