Proc. Nat. Acad. Sci. USA
Vol. 69, No. 5, pp. 1192–1195, May 1972

Low Activity of β-Galactosidase in Frameshift Mutants of *Escherichia coli*

(*lac* operon/ribosomal ambiguity/translational leakiness/acridine mutagenesis)

JOHN F. ATKINS*, DIRK ELSEVIERS, AND LUIGI GORINI†

Department of Molecular Biology, University of Aarhus, Denmark; and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115

Contributed by Luigi Gorini, March 9, 1972

ABSTRACT **16 *lac* frameshift mutants induced by an acridine derivative, ICR-191D, in *E. coli* are leaky for β-galactosidase activity. Activities of all mutants differ from each other and from the wild type in their stability to thermal denaturation. The leakiness is under ribosomal control, since it is strongly reduced by *strA* restrictive mutations and is restored by *ram* mutations that reverse restriction. Addition of streptomycin during growth has an effect similar to the presence of the *ram* mutation. These ribosomal alterations do not modify the thermal stability of the enzyme.**

It is suggested that the leakiness is due to an infrequent 2- or 4-base reading close to the frameshift mutation site. The possibility that not only the ribosome, but also the reading context in the messenger, plays a role in securing code fidelity is discussed.

Studies with frameshift mutations in phage T4 revealed the triplet nature of the genetic code (1, 2). Subsequently, frameshift mutants of bacteria have been studied to find out how fixed the present genetic code is in always having three as the reading unit. Frameshift mutations can be externally suppressed (3-5) by the reading of a non-triplet codon (4). Recent studies (6) have demonstrated that one class of external suppressors is altered in the structural gene for a minor tRNA species, and another class is defective in the enzymic modification of a minor tRNA species.

It was originally assumed to be obvious that frameshift mutants should be completely negative unless the mutation occurs toward the end of the peptide chain, where an altered aminoacid sequence might not completely damage protein activity. In one case, leakiness was interpreted as a consequence of in-phase reinitiation at the site of the frameshift mutation (3). More recently (7), however, a set of 100 *Escherichia coli* mutants, presumably frameshift because they were induced by ICR-191D, were selected for their inability to grow on lactose; 50% of them carried a defective β-galactosidase (coded for by the *z* gene of the *lac* operon). Surprisingly, several of these *z* mutants were not strictly Lac$^-$ on a minimal lactose plate (7). Only the nonleaky strains were studied further; 31 of them were mapped, and were shown to be frameshift mutants by several criteria (7).

We wanted to verify the presumed absence of β-galactosidase activity in these 31 strains by using a test more sensitive than ability to grow on lactose. 16 Mutants, scattered throughout the *z* gene (see Fig. 1), were chosen and tested on minimal glycerol plates containing 1 mg of 5-bromo-4-chloro-3-indolyl-β-D-galactoside and 0.24 mg of IPTG (isopropyl-β-D-thiogalactoside) per plate. Upon incubation at 37°, the colonies of all the 16 strains tested aquired a more- or less-intense blue color (after 1-5 incubation days), showing that all the colonies were able to hydrolyse 5-bromo-4-chloro-3-indolyl-β-D-galactoside at different rates. By contrast, no color appeared, even after 10 days, with strains CP179 (*lac* deletion X74) or CA7033 (*lac pro* deletion).

This result suggested that leakiness among frameshift mutants is more common than was thought originally. Furthermore, since no correlation is visible between color intensity and the mutant's map position, explanations based on particular locations—such as closeness to the end of the gene or to possible reinitiation sites—are unlikely. The present study is an attempt to interpret frameshift leakiness as a consequence of translational ambiguity controlled by the same ribosomal mutations known (8) to affect leakiness of nonsense mutants.

BACTERIAL STRAINS

All 16 *lac* mutants tested in this study were isolated from *E. coli* strain 3000 (Hfr Hayes), mapped and characterized by A. Newton (7). The frameshift nature of their *z* mutation was established by the following criteria: (*a*) They were induced by ICR-191D, a mutagen that causes base insertions or deletions; (*b*) 15 Were revertible to Lac$^+$ by the same mutagen, but six of them (ICR55, ICR26, ICR11, ICR38, ICR5, and ICR24) were also revertible by nitrosoguanidine (*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine), albeit at a much lower frequency. Mutant ICR17 was not revertible with either mutagen, and might be a deletion more extensive than one base; (*c*) All mutants were polar for the expression of the distal *a* gene, coding for transacetylase; (*d*) None of the mutants was suppressible by nonsense suppressors, such as amber *suI* and *suIII*, ochre *suB* and *suC*, and one for UGA; (*e*) Two mutations, *36 14*$^-$ and *36 18*$^-$, which have the above characteristics and, therefore, are frameshift, were isolated as internal suppressors of frameshift ICR36.

All strains were *strA*$^+$ except for mutations *lac 36*, *lac 36 14*$^-$, and *lac 36 18*$^-$, which were available to us on episomes carried by an *strA* mutant strain. They were mated with the *strA*$^+$ *lac-pro* 3000 XII deletion strain in a medium, with melibiose as carbon source at 42° (9). *StrA*$^+$ F'-ductants were

Abbreviations: ICR-191D, 3-chloro-7-methoxy-9-(3-[chloroethyl]aminopropyl-amino) acridine dihydrochloride; IPTG, isopropyl-β-D-thiogalactoside; Sm, streptomycin.

* Present address: Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, England.

† Reprint requests should be sent to Luigi Gorini.

isolated, and the *lac* mutation was checked by crossing with strains carrying the same mutations on the chromosome.

The *strA* alleles were introduced by P1 transduction and selection for Sm^R recombinants (10). The *ram1* allele was introduced by P1 transduction, along with the 98% cotransducible *spc1* mutation, by selection for Spc^R (10).

RESULTS

The histogram constructed on top of the *z* gene map (Fig. 1) confirms the semiquantitative indication given by the experiment with 5-bromo-4-chloro-3-indolyl-β-D-galactoside plates. Namely: (*a*) All of the 16 strains tested possess β-galactosidase activity, the difference between the two extremes (ICR38 and *36 18*$^-$) being about 100 times; (*b*) No pattern connecting the differences in activity and map location is found. It is also noteworthy that ICR36 and its internal suppressor *36 14*$^-$ possess comparable levels of activity, indicating that both classes of frameshift, base addition and base deletion, might be leaky. Some numerical values of activity are given in Table 1. The highest observed activity in the mutants is 0.06% of that of the wild type. This finding does not necessarily mean that frameshift leakiness is bound to be low, since a first screening (7) had eliminated all those mutants that were leaky on minimal lactose plates (and, therefore, possess 0.1% or more of the wild-type activity), yet they were obtained in the same selection and were, presumably, frameshift mutants.

Table 1 shows that mutations in the ribosomal gene *strA* [that codes (11) for protein P10 of the 30S ribosomal subunit] that are known (8) to restrict leakiness of nonsense mutants also restrict leakiness of frameshift mutations. Nonsense restriction is reversed by mutations in *ram* (10),

TABLE 1. *β-Galactosidase activity in ICR mutants*

ICR mutant	β-Galactosidase activity in the following ribosomal context, %		
	strA$^+$ *ram*$^+$	*strA1* *ram*$^+$	*strA1* *ram1*
5	100	20	116
24	85	29	—
38	98	44	93
55	109	29	132
48	50	19	—
26	17	2	—
75	16	2	26
11	0.6	—	—
21	5.6	—	—

The cells were grown in 5 ml of medium, as in Fig. 1; when the culture reached a density of 4×10^8 cells/ml, growth was stopped by cooling in ice, followed immediately by determination (20) of absorbance and toluenization. β-Galactosidase determination was performed directly on the toluenized suspensions, as described in Fig. 1. Specific activity is expressed per dry weight of cells. Data are the average of triplicate determinations; experimental error was within 10%, except for values lower than 20. The relative activities of the different ICR mutants shown are slightly different from those shown in the histogram of Fig. 1, presumably due to the fact that the mutant enzymes have different stabilities, and toluenized cells (instead of extracts obtained by sonication) were used for the determinations in the tables. The activity in strain CA7033 (*lac pro*)$^\Delta$, tested as a control, is 0(or <0.2%) of that in ICR5. The value of ICR5 activity, set at 100, is in fact 0.05% of the wild-type activity.

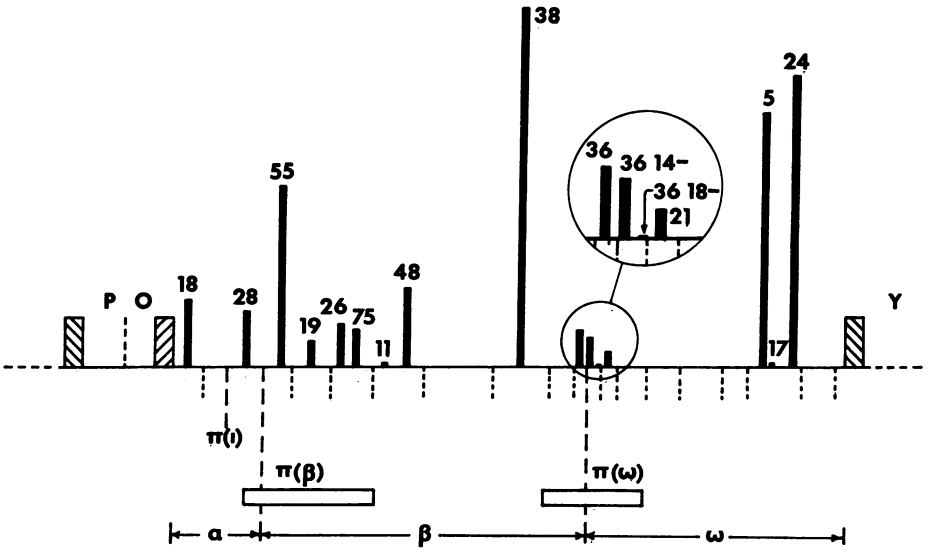

FIG. 1. Map position of *lac z* frameshift mutations (7) correlated with positions of internal restart regions (13–15) and with levels of β-galactosidase. The relative degrees of leakiness are shown by the *histogram* above the *horizontal line*. The *dotted lines* represent the end-points of the deletions used for mapping mutation order. The order of mutations between adjacent deletion end-points has not been determined. π (β) and π (ω) are natural internal restart regions; π (*1*) is a position where an initiator sequence is readily produced by mutation. α, β, and ω are complementation regions. For determination of enzyme activity, the cells were inoculated to a density of 10^8 cells/ml into 200 ml of minimal medium A (16) containing 1 μg of thiamine/ml, 1 mM IPTG, and 0.02% glycerol as carbon source. The inoculum was an overnight culture in the same medium. Incubation was at 37° with aeration. When a density of 4×10^8 cells/ml was reached, the cells were harvested, washed, frozen, and sonicated, as described by Ullman *et al.* (17). The β-galactosidase assay procedure was that of Epstein (18). Assay for very-low activities required incubation for 24–48 hr. This prolonged incubation involves a loss of about 20% of the activity over a short-term incubation, as was shown by comparison of values obtained for wild-type cells in an undiluted sample incubated 1 min and in a sample diluted 1:5000 and incubated 48 hr. Therefore, both an undiluted sample and a sample of wild-type cells diluted 1:5000 were run as reference in every assay. Protein was determined according to Lowry *et al.* (19).

TABLE 2. *Influence of ribosomal mutations and of the presence of Sm on leakiness of ICR5*

Ribosomal genotype		Presence of Sm	β-Galactosidase activity (% of *strA*$^+$ *ram*$^+$)
strA	*ram*		
+	+	No	100
40	+	No	48
40	+	Yes	86
41	+	No	29
41	+	Yes	57
11	+	No	9
11	+	Yes	51
1	+	No	24
1	+	Yes	26
1	1*	No	120

Alleles *strA40* and *strA1* were isolated from strain B as *strA* mutations that restrict leakiness of amber mutant *argF40*. Restriction by *strA40* is reversed phenotypically by Sm, while that by *strA1* is irreversible. Alleles *strA41* and *strA11* were similarly isolated from strain K. They restrict leakiness of *his41*, an amber mutation in the *his* operon. Restriction by *strA41* is reversible by streptomycin, that by *strA11* is not. Allele *ram1* was isolated from strain B as a ribosomal mutation able to reverse restriction by *strA1* of amber *argF40*. Sm is added to the growth medium at a concentration of 500 μg/ml. The same Sm concentration added to toluenized cells does not affect the determination of β-galactosidase activity. Data are averages of triplicate determinations; experimental error is within 10%, except for values lower than 10, in which case it is 30%.

* This *ram1* derivative carries the mutation *spc1*, introduced along with *ram1* by P1-transduction. This mutation is irrelevant to leakiness, since strains *strA1 ram*$^+$ *spc*$^+$ and *strA1 ram*$^+$ *spc1* possess comparable levels of activity (24 and 22, respectively).

another 30S ribosomal gene that codes for protein P4a (12), and in some cases by growth in the presence of Sm (streptomycin) (8). Differences in the degree of restriction and in the reversion patterns of these mutants have allowed a classification of *strA* alleles. Tables 1 and 2 show that this classification is remarkably similar when it is based on leakiness of frameshift mutations.

The heat stability of the β-galactosidase activity, contained in the different frameshift mutants, and in their *strA* and *strA ram* derivatives, has been determined as a test of the structural similarity of the enzymes. Table 3 indicates (*a*) that in no case is the enzyme produced by the mutants identical to that of the wild type; and (*b*) that the ribosomal mutations restricting and derestricting leakiness do not introduce alterations in the heat stability that are characteristic for a given frameshift mutant.

DISCUSSION

We have found that low enzyme activity (leakiness) occurs frequently among frameshift mutants of the β-galactosidase structural gene, and does not appear to be related to any particular map location, either at the end of the gene or at the end of any of the complementation regions into which this gene is subdivided. We find, moreover, that this leakiness is controlled by ribosomal alterations. It is, therefore, excluded that this type of leakiness might be due to faulty, but not completely inactive, β-galactosidase protein produced through continued translation in the altered reading frame. The possibility that an alternative path might replace, albeit poorly, the function of the inactive β-galactosidase is ruled out by the absence of activity in the *lac* deletions We conclude, instead, that this type of leakiness is due to the existence of alternate translation possibilities able to suppress, at a low level, the effect of the frameshift mutation.

The ribosomal alterations that control frameshift leakiness are the same *strA* and *ram* mutations and Sm-induced phenotypic alterations in the 30S subunit that are known to control nonsense leakiness. The difference in restriction by different *strA* alleles and their reversibility, by *ram* or by Sm, are remarkably parallel in the case of frameshift and of nonsense leakiness. Thus, the indication is strong that the mechanism underlying leakiness of both types of mutations might be the same. In the case of nonsense mutations, leakiness has been explained as due to a small amount of translation performed by normal tRNA, not specific for nonsense codons (8). In the case of frameshift mutations, it might be suggested that normal, nonmutated tRNAs could read 2- or 4-letter codons at a low frequency. Such a mechanism is known to be used by specific frameshift suppressors that appear to be mutated or modified tRNAs (3-6).

However, if the genetic code is to serve its purpose, possibilities for ambiguity should be very limited. The occurrence of ambiguous reading and its limitation is easy to accept in the case of nonsense codons, because weakly-binding nonspecific tRNAs are not in competition with any strongly-binding, nonsense-specific tRNA and, moreover, the occurrence of a nonsense codon within a cistron is an abnor-

TABLE 3. *Percent of activity remaining after 5 min at 50°*

Ribosomal modifications of ICR5			Other ICR mutants (*strA*$^+$ *ram*$^+$)		
	β-Galactosidase activity			β-Galacotsidase activity	
ICR5 derivatives	Initial	Residual %	ICR mutant	Initial	Residual (%)
strA$^+$ *ram*$^+$	(100)	46	24	(85)	42
strA1 ram$^+$	(24)	47	38	(98)	69
strA1 ram1	(120)	50	55	(109)	97
strA11 ram$^+$	(9)	54	48	(50)	76
strA11 ram$^+$ + Sm*	(51)	48	26	(17)	88
			75	(16)	100

Heat inactivation was performed in a water bath at 50° by incubation of toluenized cells suspended in their culture fluid at comparable densities. Samples were withdrawn after 5, 10, and 20 min, and β-galactosidase activity was determined. The inactivation time reported in Table 3 (5 min) is that during which the steepest drop in activity is observed. Wild-type enzyme is stable at 50°, even after 20 min. Enzyme from ICR55 is less stable than the wild type, since after 5 min at 57° the residual activity is, respectively, 20 and 100% of the initial activity. Enzyme from ICR75 is slightly less stable than the wild type when heated at 60°. The values of initial activities (given within parentheses) are in percent of that of the ICR5 *strA*$^+$ *ram*$^+$ strain.

* 500 μg/ml of growth medium.

mality. The situation, however, is quite different with frameshift mutations, because an out-of-phase codon appears normal as far as availability of a specific tRNA is concerned.

Two possibilities might be considered: (*a*) The ambiguous reading of 2 or 4 bases as a single codon could occur at the first nonsense codon generated by the frameshift. This could happen, as the sequence surrounding such a nonsense triplet does not correspond to the neighboring sequences around natural terminators. This hypothesis implies a sequence of wrong amino acids between the start of the frameshift and the first nonsense codon. On the average, the length of such a sequence is 20 amino acids. This should strongly limit the occurrence of leakiness, but such does not seem to be the case for the β-galactosidase frameshift mutants; (*b*) The site at which the addition or deletion of a base occurs might not be as comparable to any other site as it superficially looks; this could be the site of suppression. Apparently, frameshift mutations occur predominantly in DNA sequences of repeated bases, and consist of the addition or deletion of a base identical to the adjacent ones (4, 21-24). It may be that a given tRNA, faced with a monotonous sequence of identical bases, encounters difficulty in maintaining a particular phase because its anticodon can pair equally well in two or three successive phases. Hypothesis (*b*) implies that leakiness is due to a few molecules of enzyme protein identical (or as close as one aminoacid difference) to the wild-type enzyme. Amino-acid sequence analysis of a few active enzyme molecules occurring in traces in the frameshift translation product probably is not feasible. However, as a first approximation, heat stability of the enzyme activity responsible for leakiness has been compared with that of the wild-type enzyme. We anticipated that finding enzymes with identical stability should support hypothesis (*b*). It was found, however, that all the enzymes responsible for the leakiness of each frameshift mutant are different among themselves, and different also from the wild type.

Other mechanisms have been suggested to explain frameshift suppression. One involves termination and reinitiation of translation. In the case of leakiness, this hypothesis should imply that (*a*) the frameshift mutation is located near the end of a complementation region; (*b*) an out-of-frame nonsense codon occurs before the initiation point of the next complementation region; and (*c*) the reinitiation is controlled by the *strA*, *ram*, and Sm ribosomal modifications. All these implications are contradicted by the random location of leaky frameshift mutations, and by the lack of evidence that the ribosome controls translation specifically at the initiation points. This last argument concerning ribosomal control also seems to invalidate any hypothesis attempting to explain frameshift leakiness by postulation of a ribosomal slippage along the message, since there is no evidence that the same factors (*strA*, *ram*, and Sm) controlling tRNA translation efficiency should also control a general slippage of the ribosome.

In conclusion, some ambiguity, intrinsic to the genetic code and controlled by ribosomes, seems to be at the origin of frameshift leakiness and leakiness of nonsense mutations. It is seen in both cases that the amount of ambiguity allowed by the wild-type ribosome is higher than that allowed by *strA*-mutated ribosomes (25). Since it is found that *strA* mutations also restrict translation efficiency of at least some tRNAs (26), it might be suggested that through natural selection the translation mechanism ended up with some ambiguity as a price to pay for speed. The identical heat stability of the *ram*$^+$ and *ram1* derivatives of the same frameshift mutant indicates that these ribosomal mutations (or ribosomal modification by Sm) do not introduce new structural differences in the enzyme molecule responsible for leakiness of any given frameshift mutant. This finding confirms previous results, obtained with nonsense mutants, indicating that *ram* or Sm addition increase the amount of ambiguity, but do not produce new types of misreading (8). Finally, it appears that the proposed mechanism for frameshift leakiness is acceptable only if a recognition site for ambiguous reading can be postulated in the messenger [hypotheses (*a*) or (*b*) above]. The influence of the reading context already found in nonsense suppression (20, 27) is, thus, confirmed.

We thank Dr. A. Newton for the generous gift of the collection of mutants that made this project possible and Dr. H. J. Creech, Institute for Cancer Research, for the supply of ICR191. J.F.A. thanks Dr. N. O. Kjeldgaard for his advice and the hospitality of his laboratory, and acknowledges receipt of an EMBO fellowship for part of his work. The portion of this work done at Harvard was supported by United States Public Health Service Grant 5 R01 AI 02011-15, National Institute of Allergy and Infectious Diseases, and by American Cancer Society Grant VC-8J. D.E. is an "Aspirant du Fond National de la Recherche Scientifique" of Belgium.

1. Crick, F. H. C., Barnett, L., Brenner, S. & Watts-Tobin, R. J. (1961) *Nature* **192**, 1227–1232.
2. Terzaghi, E., Okada, Y., Streisinger, G., Emrich, J., Inouye, M. & Tsugita, A. (1966) *Proc. Nat. Acad. Sci. USA* **56**, 500–507.
3. Riyasaty, S. & Atkins, J. F. (1968) *J. Mol. Biol.* **34**, 541–557.
4. Yourno, J. & Tanemura, S. (1970) *Nature* **225**, 422–426.
5. Riddle, D. L. & Roth, J. R. (1970) *J. Mol. Biol.* **54**, 131–144.
6. Riddle, D. L. & Roth, J. R. (1971) *Genetics Suppl.*, S54.
7. Newton, A. (1970) *J. Mol. Biol.* **49**, 589–601.
8. Gorini, L. (1972) *Nature New Biol.* **234**, 261–264.
9. Beckwith, J. R. (1963) *Biochim. Biophys. Acta* **76**, 162–164.
10. Rosset, R. & Gorini, L. (1969) *J. Mol. Biol.* **39**, 95–112.
11. Ozaki, M., Mizushima, S. & Nomura, M. (1969) *Nature* **222**, 333–339.
12. Zimmermann, R. A., Garvin, R. T. & Gorini, L. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 2263–2267.
13. Grodzicker, T. & Zipser, D. (1968) *J. Mol. Biol.* **38**, 305–314.
14. Michels, C. A. & Zipser, D. (1969) *J. Mol. Biol.* **41**, 341–347.
15. Newton, A. (1969) *J. Mol. Biol.* **41**, 329–339.
16. Davis, B. D. & Mingioli, E. S. (1950) *J. Bacteriol.* **60**, 17–28.
17. Ullmann, A., Perrin, D., Jacob, F. & Monod, J. (1965) *J. Mol. Biol.* **12**, 918–923.
18. Epstein, W. (1967) *J. Mol. Biol.* **30**, 529–543.
19. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275.
20. Strigini, P. & Gorini, L. (1970) *J. Mol. Biol.* **47**, 517–530.
21. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 77–84.
22. Yourno, J. & Heath, S. (1969) *J. Bacteriol.* **100**, 460–468.
23. Tanemura, S. & Yourno, J. (1969) *J. Mol. Biol.* **46**, 459–466.
24. Yourno, J. & Kohno, T. (1972) *Science* **175**, 650–652.
25. Gorini, L. (1969) *Cold Spring Harbor Symp. Quant. Biol.* **34**, 101–109.
26. Biswas, D. K. & Gorini, L. (1972) *J. Mol. Biol.* **64**, 119–134.
27. Salser, W. (1969) *Mol. Gen. Genet.* **105**, 125–130.