J Mol Evol (1995) 40:7–12

JOURNAL OF MOLECULAR EVOLUTION

The Generation of Variation in Bacterial Genomes

W. Arber

Department of Microbiology, Biozentrum, University of Basel, Klingelbergstr. 70, CH-4056 Basel, Switzerland

Received: 5 October 1994

Abstract. In the context of a general overview of molecular mechanisms of microbial evolution, several genetic systems known to either promote or restrain the generation of genetic variations are discussed. Particular attention is given to functions involved in DNA rearrangements and DNA acquisition. Sporadic actions by a variety of such systems influencing genetic stability in either way result in a level of genetic plasticity which is tolerable to the overall wealth of microbial populations but which allows for evolutionary change needed for a steady adaptation to variable selective forces. Although these evolutionarily relevant biological functions are encoded by the genome of each individual, their actions are exerted to some degree randomly in rare individuals and are therefore seemingly nondeterministic and become manifest at the population level.

Key words: DNA acquisition — DNA rearrangement — Evolutionary function — Genetic variation — Genomic plasticity — Illegitimate recombination — Selection — Site-specific recombination — Spontaneous mutagenesis — Transposition — Variation generator

Introduction

Whether biological functions—and thus the products of genetic determinants—are specifically involved in processes of biological evolution remains an open question. At least partial answers can be expected to be found in molecular genetic experimentation and analysis, although the interpretation of available data may often depend on the conceptual attitude that the investigator has toward nature, life, and its evolutionary development. It is in this context that knowledge accumulated on genetic plasticity will be reviewed with particular attention given to molecular mechanisms of genetic alterations as revealed in work with haploid microorganisms.

Conventional Views on the Generation of Genetic Variation

Classical textbook knowledge defines three sources of genetic variation. First, spontaneous mutations are thought to occur randomly and in general to result from errors, mistakes, accidents, lesions. These can either originate upon DNA replication or be due to the action of internal or external environmental mutagens. Accumulation of such mutations in the course of time provides evidence for divergent evolution, and it can explain the development of altered and novel biological functions. Second, in sexual, diploid organisms the mixing of gametes of two different origins as well as homologous recombination at the molecular level between the slightly different genetic information carried in the two sets of chromosomes brings about a large variety in the individual genome composition. Fertilization by mixing of gametes is limited to organisms of the same species. Third, occasional illegitimate recombination between DNA sequences carried at different chromosomal locations can cause new and largely stochastic genome rearrangements. Only a few of these DNA rearrangements may by chance confer improved fitness to the organism; most of them are neutral or provide either selective disadvantage or lethality. Mobile genetic elements have been identified as one possible source of such illegitimate recombination, and they were thus often considered genetic parasites.

Genetic Studies with Haploid Microorganisms

Bacteria are haploid and do not depend on sexual reproduction for their propagation. There is thus no compulsory periodic mixing of genetic information carried in populations. However, several well-studied mechanisms exist which procure means for sporadic gene transfer from one bacterial cell to another. First, in transformation, DNA liberated into the environment by excretion or cell lysis can be taken up by a competent recipient cell, in which the invading genetic information may find an opportunity to become incorporated into the bacterial genome, often by substitution of related chromosomal segments. In this integration, either homologous recombination or other recombination processes can be involved. Similarly, DNA molecules equipped for autonomous replication, such as plasmids, can transform competent recipient cells in which they might find an opportunity to replicate by their own means without incorporation into a chromosome. Second, in conjugation, two genetically distinct cells can undergo a mating in which DNA from the donor becomes transferred into the recipient cell. Again, stable association of the transferred genes with the recipient genome depends on the activities of appropriate recombination systems, unless the transferred DNA is itself an autonomous replicon such as a plasmid. Third, bacterial viruses may serve as vectors for genes to be transferred from a donor to a recipient cell. The lasting effect of this process, known as phage-mediated transduction, also depends, as in conjugation and transformation, on either recombination or autonomous replication of the transferred genetic information. Finally, some bacteria have been shown to occasionally undergo cell fusion, which can result in a mixing of the genetic information of the cells involved. In conclusion, natural mechanisms exist which bring about gene transfer between individual bacterial cells. The range of such transfer is generally not restricted to a given species and can involve also distantly related organisms.

Gene Acquisition as an Evolutionary Strategy

The processes of horizontal gene transfer described above can lead to the acquisition of genetic information from other bacteria. Gene acquisition can obviously represent an interesting step in the long-term evolutionary development of bacterial strains. However, it should not occur too frequently in order to respect the need for a reasonable degree of genetic stability, which is generally thought to be a necessity to insure endurance of a species of organisms in ecosystems. A number of natural limits are indeed set to gene acquisition. These limits include the need for cell-surface compatibility of the recipient with the transfer process, the activities of restriction-modification systems, the already-discussed requirement of stable association of acquired genetic information with the recipient genome, and finally the action of natural selection on the resulting hybrid organism. For these reasons, the frequency of successful gene acquisition is quite low, and it generally becomes even lower with increasing genetic distances between the donor and the recipient organisms. Nevertheless, the evolutionary potential of gene acquisition is considerable and has been demonstrated by the efficient horizontal spread of drug resistance genes that was revealed after selection conditions for enterobacteria were drastically altered by the worldwide use of antibiotics.

Molecular Mechanisms of Generation of Genetic Variations

Experimental investigations carried out in the last 50 years, particularly with microorganisms, have unravelled a number of molecular mechanisms which contribute, each in its specific way, to the overall generation of genetic variations. Quite often, specific enzymes or enzyme systems are involved and at least some of these enzymes do not seem to assume other functions, other than to either generate genetic diversity or limit the generation of genetic variation. These enzymes are encoded by genes that form part of the microbial genomes. Therefore and for conceptual reasons, I advise not considering variations resulting from such enzyme-mediated processes as accidental. Rather, one should be ready to accept that some biological reactions catalyzed by enzymes may result in not strictly reproducible products. Examples for such seemingly nondeterministic reactions will be discussed below.

On previous occasions (Arber 1990, 1991, 1993) I have grouped the various processes generating genetic variations into four categories. The first of them consists of those processes by which mutagenesis depends on DNA replication. The actions of both external as well as internal environmental mutagens form a second category. Primary effects of mechanisms of these two categories can be reversed with considerable efficiency by enzymatic repair processes. A third category consists of various, often enzyme-mediated, DNA rearrangement processes, examples for which will be more thoroughly discussed below. A fourth category is formed by the processes of DNA acquisition that have been defined above. While mechanisms of categories 1 to 3 act on the genome of the concerned individual as substrate, DNA acquisition makes use of a foreign source of genetic information, which is either added to the recipient genome or which serves for the substitution of part of the genetic information carried by the recipient. It may be helpful to recall that different molecular mutagenesis mechanisms can result in qualitatively very different types of mutations, such as nucleotide substitutions, frame shifts, or DNA rearrangements. All of these have their specific roles in the evolutionary process. Therefore, different

mutagenesis processes should not be considered as alternative possibilities. Rather a parallel action of various mechanisms of mutagenesis is essential for the evolutionary process.

Considering the structural flexibility known to be characteristic of organic molecules, as exemplified by tautomeric forms of nucleotides which had been postulated to be a source of mutagenesis upon nucleic acid replication, one can expect structural variations to generally have an impact on the effects of molecular interactions. In this light it appears possible that the specific products of the interaction between particular enzymes with DNA as a substrate might depend on the particular, sometimes short-lived structural features displayed at the time of interaction by either the enzyme involved or its substrate DNA. We therefore postulate that both structural and intrinsic functional parameters of participants in molecular interactions involving DNA can be a general source of genetic variation. This can apply both to mutagenesis upon DNA duplication and to illegitimate, not strictly reproducible DNA rearrangements. Resulting genetic variants should thus not be looked at as aberrant, but rather as products of complex natural processes.

Mutagenesis in Resting Bacteria

Studies on spontaneous mutagenesis have often been carried out under good growth conditions and with short DNA segments such as single genes or operons. Investigations at the level of the bacterial genome and for extreme life conditions have become possible largely thanks to advances in molecular genetic technology. In a recent study we documented events of DNA rearrangements at the chromosomal level for resting bacteria of *Escherichia coli.* A sealed stab culture of strain W3110 kept at room temperature for 30 years served as a source of individual bacterial subclones. DNA of subclones was extracted and cleaved with restriction enzyme *Eco*RV. The resulting DNA fragments were separated by gel electrophoresis. Fragments containing specific DNA sequences were then identified by Southern hybridization. While hybridization with sequences from unique chromosomal genes revealed no or very little variation, a high degree of variation was detected in this analysis of restriction-fragment-length polymorphism (RFLP) if internal sequences of mobile genetic elements (IS elements) resident in strain W3110 were used as hybridization probes (Naas et al. 1994). Evidence was thereby obtained that many of these variations must have been caused by transpositional recombination.

Subsequent studies with stab cultures of different ages indicated a linear increase in the degree of polymorphism in bacterial populations with time of storage (T. Naas, M. Blot, W. Arber, to be published). This must be correlated with the steady availability of transposase in resting bacteria of stab cultures, which at most can allow for a very limited growth on the residue of dead cells.

These results confirm and extend previous conclusions to the effect that resting cells in stab cultures can be affected by mutagenesis and that transposition is an important cause of such mutagenesis (Arber et al. 1979; Green et al. 1984). The possibility of correlating individual subclones of strain W3110, even of different stabs, with regard to their common descent by the construction of parsimonious pedigrees on the basis of their RFLP pattern supports the conclusions on the evolutionary relevance of transpositional events as contribution to spontaneous mutagenesis (Naas et al. 1994; W. Fitch, personal communication).

Interplay Between Mutation and Selection

It is generally agreed that newly occurring mutations only rarely provide a selective advantage to the organism undergoing a sequence alteration. More often mutations are neutral or they procure a selective disadvantage or even lethality. On the basis of these considerations one postulates for haploid organisms that rates of overall spontaneous mutagenesis should be kept below one mutation per genome and per generation. For *E. coli* bacteria kept under standard growth conditions this rate is between 1% and 0.1%. However, in the absence of exponential growth, such as in stab cultures, it is difficult, if not impossible, to define the generation time. On the other hand, one can assume that the effects of selection are less strong in populations of resting bacteria than in exponentially growing cultures. Therefore, even less fit bacterial mutants can persist for long periods of time in stabs. This expectation was indeed verified in an analysis of competition between subclones grown from old stab cultures (Naas et al. 1994; J. Meier, M. Blot, and W. Arber, unpublished results).

The fact that cultures of *E. coli* withstand long periods of starvation in stabs indicates that these bacteria are intrinsically able to undergo this type of "hibernation," which of course can also be considered a particular kind of selective force. Since bacteria may often encounter phases of starvation under natural conditions, it is not surprising that they adapted, in the course of their evolution, means to cope with such drastic extremes in their life conditions. In this light it is obvious that natural selection should not be considered to be exerted only under optimal growth conditions and that bacterial fitness is not limited to efficient growth. Rather, it is also of primary importance for bacteria not to die in their resting phases. Therefore the steady interplay between selection and mutation is a complex phenomenon and may vary in its specificity with time and space. In this regard it is interesting that under so-called adverse conditions of life, enzyme-mediated mutagenesis (by transposition) is maintained as one of the physiological activities of resting bacteria.

Maintenance of Tolerable Levels of Genome Plasticity

Because of changing needs to genetically adapt to variable living conditions, a certain degree of genetic instability is beneficial to microbial populations. But the genome plasticity should not be excessive in view of the often-adverse effects of spontaneous mutagenesis to fitness, as we have already discussed. Upon close examination of the situation, one often sees a kind of equilibrium established between effects promoting genetic variation and effects limiting genetic variation. For example, mutagenesis by tautomerism of nucleotides is counteracted by efficient repair systems. Similarly, gene acquisition is promoted by active systems of horizontal gene transfer, but it is limited by surface incompatibilities, by restriction systems, and by selective forces exerted at the level of the biological functions expressed by the resulting hybrid cells.

As we have already seen, transposition is an important source of genetic plasticity under various physiological conditions. For several well-studied mobile genetic elements, transposition frequencies are quite low, except for bacteriophage Mu in its cycle of vegetative growth (Mizuuchi and Craigie 1986). Interestingly, in bacterial IS elements the expression of transposase is normally initiated by promoters which do not show excessively low strength. However, various factors limiting transposition efficiencies have been described. A few examples should illustrate this statement. By their specific binding to IS ends, transposases may repress their own synthesis, if their binding site overlaps with their transcription promoter, as exemplified with IS*1* (Zerbib et al. 1990). In IS*30,* an internal leaky transcription terminator allows only a minority of RNA transcripts to attain their full length, required for the functional transposase (Dalrymple and Arber 1986; Olasz et al. 1993). At the translational level, limits are set to the transposition activity of different transposable elements by an unconventional codon usage for transposase expression (Médigue et al. 1991). In addition, IS*10* and IS*30* have been shown to constitutively produce antisense RNA transcripts which inhibit an efficient production of transposase (Simons and Kleckner 1983; Arini et al., to be published). Some IS elements such as IS*1* require a translational frameshift to occur in order to produce an active transposase (Sekine and Ohtsubo 1989). At the posttranslational level, target selection criteria may considerably limit transposition efficiencies.

Sequence Specificity vs Randomness of Mutagenesis

Transposition target specificity can ideally serve in a discussion of the sequence dependence of mutagenesis. Each transposable element shows its specific characteristic behavior with regard to target selection. Some IS elements show a high degree of target specificity, but they may also use other insertion sites, although with much lower frequencies than the preferred target. Other IS elements have been thought to select their insertion targets largely at random. But a closer analysis has revealed in some cases that this may not apply at the genome level.

A striking illustration of these general statements was seen in a study carried out with a P1-lysogenic derivative of *E. coli* K-12. In P1-lysogens the P1 prophage is maintained as a plasmid. In the prophage state spontaneous mutations of phage functions can accumulate, even if they are potentially lethal for vegetative phage reproduction. Such lethal mutations have been screened for after prolonged periods of interrupted growth of a culture of lysogens for an IS*1*-free derivative of P1. (Wild-type P1 carries an IS*1* element in its genome.) A molecular genetic analysis of the mutant prophages revealed that about 95% of independent phage mutants carried an IS element that must have transposed from the host chromosome to the prophage, while the nonmutated prophage in the studied population had not suffered an IS transposition. Several different IS elements have contributed to this mutagenesis (Arber et al. 1981; Sengstag and Arber 1983). At the genome level, the IS insertions had not occurred at random. Most strikingly, a 1.8-kb segment of the 90-kb P1 genome contained more than half of all detected IS*2* insertions as well as all three independently obtained IS*30* insertions. Interestingly, a detailed analysis of specific insertion sites revealed all three IS*30* insertions to be exactly at the same site, and the element was found in both possible orientations. In contrast, all of the nine sequenced IS*2* integration sites in this region were unique, and their distribution on the 1.8-kb genome segment appeared to be almost random. Hence, IS*2* shows a regional specificity of target selection but it can integrate at many different and nonhomologous sequences within the preferred DNA segment. This conclusion was later confirmed by subcloning specific DNA segments from the P1 genome region preferred by IS*2* for transposition. Evidence was thereby obtained that target selection is nevertheless not fully statistical. Rather, an occasionally repeated use of some given target sites was observed (Sengstag and Arber 1987).

Genetically Determined Variation Generators

The relaxed target specificity of IS*2* is a good example for an enzyme-mediated generator of genetic variation. In transposition the IS*2* sequences are specifically translocated to a target site, but one of a large number of potential target sites may be chosen. Therefore, the specific product of a single transposition event is not predictable, and the reaction is thus not strictly reproducible.

This kind of nonreproducible interaction of enzymes with DNA is seen also in a number of other cases of DNA–protein interactions, two of which shall be presented here. It is the function of restriction enzymes of type I to specifically recognize DNA molecules of foreign origin upon their penetration by horizontal gene transfer into recipient bacteria. Although the recognition occurs at specific DNA sequences, generally on the basis of their lack of sequence-specific DNA methylation, the subsequent fragmentation of foreign DNA molecules occurs relatively randomly by cleavage at variable distances from the specific recognition sites (Studier and Bandyopadhyah 1988). Since DNA fragments with free ends are recombinogenic, this strategy of cleavage of restriction enzymes may achieve its evolutionary relevance by stimulating the acquisition of small portions of genetic information of various content, while drastically reducing the overall probability of DNA acquisition upon horizontal DNA transfer.

Another example of a variation generator is DNA inversion by site-specific recombination. Such DNA inversion systems have been well studied both in vivo and in vitro. In the presence of two 26-bp consensus crossover sites at some distance on the same DNA molecule and in inverted orientations, the availability of DNA invertase in the same cell and the presence of a site for inversion stimulation (also called recombinational enhancer) on the same DNA molecule as the crossover sites bring about a periodic inversion of the DNA segment carried between the crossover sites. This kind of periodic DNA inversion is known as a flip-flop system (Glasgow et al. 1989). In contrast, if a DNA molecule carries only one consensus crossover site and a recombinational enhancer, no efficient DNA inversion takes place, but at very low frequencies DNA invertase can bring about DNA inversion by involving, besides the consensus crossover site, one of many different so-called secondary or quasi-crossover sites. Such secondary recombination sites have been sequenced and these sequences may deviate extensively from those of the much more efficiently used consensus crossover sites. It appears that each secondary crossover site has its own specific probability to serve for DNA inversion, probably as a function of its specific structural features (Arber 1990, 1991; Iida and Hiestand-Nauer 1986, 1987; Iida et al. 1984).

In recent experiments it has been shown that secondary crossover sites are not only used for DNA inversion on plasmids but also on the chromosome of *E. coli* and that such inversions can span distances of more than 200 kb (Rozsa et al., to be published).

Rare, aleatoric DNA inversion provides the evolutionarily interesting potential to occasionally bring about novel gene fusions (reading frames composed of functional domains and sequence motifs of different origin) and operon fusions (genes brought under different expression control). Only a single step of recombination is required in this process and this is not coupled with any loss of DNA sequences. Often, the recombination may reduce fitness or cause lethality. However, since only rare cells in large populations undergo such a DNA rearrangement, the process cannot do harm to the population. Rather, it is the rare successful DNA rearrangements resulting in increased fitness which count. On the basis of these considerations we have postulated this kind of enzymatic DNA inversion systems to primarily serve biological evolution by their rarely occurring recombination activity involving secondary crossover sites (Arber 1991, 1993). If these enzymes also drive a flip-flop system between two or more consensus sites, selection acting on the population will decide if such activity is tolerable. In the best case, a flip-flop system may even be of some use. But one may wonder if driving of a flip-flop mechanism alone could justify the presence of a DNA inversion system. Its important biological function could rather reside in its potential to act as a variation generator, an evolutionary function carried out at the population level. Support for this hypothesis is found in the structural arrangement of DNA sequence motifs common to the tail-fiber genes of several otherwise-unrelated bacteriophages (Sandmeier 1994). Interestingly, this kind of evolutionary function is not directly a target for its selection. Rather its selection at the population level (group selection) occurs by its occasional mutagenic production of more fit variants of any kind of directly selected traits.

Conceptual Considerations: Parasitic Elements or Genes with Evolutionary Functions?

The philosophy expressed in this article—that genetic variation is largely a function of partly aleatoric activities of enzymes with evolutionary roles and that genetic variation also depends to some degree on the structural flexibility of organic compounds—contrasts with the more-classical view of biochemistry: that enzyme-catalyzed reactions reproducibly result in predictable products of uniform characteristics. This latter attitude anchors the concept that spontaneous mutations are attributable to errors, mistakes, and accidents. The postulate that mobile genetic elements are parasites of the genome and act selfishly belongs to the same line of thought.

I believe that neither of these two extreme views fully corresponds to reality. Rather I see them as reflecting, together, a deeply dualistic complex situation. It is true that many spontaneous alterations of the genetic information do affect the life quality of the individual suffering the mutation, and rather rarely does a mutation improve fitness. However, since it becomes more and more clear that very often specific enzymes and enzyme systems are involved in spontaneous mutagenesis (including partly aleatoric DNA arrangements), the time has come

to admit that there are genes carried in microbial genomes (and possibly in other genomes as well) with primarily evolutionary functions. Many of these functions are not strictly deterministic, and their action is often detrimental. But at the population level this is counterbalanced by various factors keeping the activities of variation generators low. Among these factors are enzyme systems which reduce the mutagenic efficiencies and thus also exert evolutionary functions. Members of this latter class of enzymes are, e.g., DNA repair systems and restriction-modification systems. Because of their involvement in the DNA acquisition strategy, one may also consider viral genomes and conjugative plasmids as belonging to the class of genetic elements with primarily evolutionary functions. However, a strict separation between, on the one hand, genes serving the needs of individual lives (housekeeping genes and genes with accessory functions for use by all individuals under particular living conditions) and, on the other hand, evolutionary genes may not be fully justified. Indeed, a biological function expressed in a cell may carry out its activities whenever and wherever it finds an opportunity to do so. Good examples of functions required both for the successful individual lives and for the evolution of populations are topoisomerases and DNA ligases. Still, we believe that a conceptual definition of genetically encoded evolutionary functions with often-nondeterministic activities—and therefore without an implementation to drive evolution in a specific direction—may represent an important step toward a better understanding of the process of biological evolution.

Acknowledgments. The author thanks Charles Thompson and Frank Rozsa for their editorial help. The continuous research support received from the Swiss National Science Foundation (grants 31.30040-90 and 5001-35228) is acknowledged.

References

Arber W (1990) Mechanisms in microbial evolution. J Struct Biol 104:107–111

Arber W (1991) Elements in microbial evolution. J Mol Evol 33:4–12

Arber W (1993) Evolution of prokaryotic genomes. Gene 135:49–56

Arber W, Hümbelin M, Caspers P, Reif HJ, Iida S, Meyer J (1981) Spontaneous mutations in the *Escherichia coli* prophage P1 and IS-mediated processes. Cold Spring Harbor Symp Quant Biol 45: 38–40

Arber W, Iida S, Jütte H, Caspers P, Meyer J, Hänni C (1979) Rearrangements of genetic material in *Escherichia coli* as observed on the bacteriophage P1 plasmid. Cold Spring Harbor Symp Quant Biol 43:1197–1207

Dalrymple B, Arber W (1986) The characterization of terminators of RNA transcription on IS*30* and an analysis of their role in IS element-mediated polarity. Gene 44:1–10

Glasgow AC, Hughes KT, Simon MI (1989) Bacterial DNA inversion systems. In: Berg DE, Howe MM (eds) Mobile DNA. American Society of Microbiology, Washington, DC, pp 637–659

Green L, Miller R, Dykhuizen D, Hartl D (1984) Distribution of DNA insertion elements IS*5* in natural isolates of *Escherichia coli.* Proc Natl Acad Sci USA 81:4500–4504

Iida S, Hiestand-Nauer R (1986) Localized conversion at the crossover sequences in the site-specific DNA inversion system of bacteriophage P1. Cell 45:71–79

Iida S, Hiestand-Nauer R (1987) Role of the central dinucleotide at the crossover sites for the selection of quasi sites in DNA inversion mediated by the site-specific Cin recombinase of phage P1. Mol Gen Genet 208:464–468

Iida S, Huber H, Hiestand-Nauer R, Meyer J, Bickle TA, Arber W (1984) The bacteriophage P1 site-specific recombinase Cin: recombination events and DNA recognition sequences. Cold Spring Harbor Symp Quant Biol 49:769–777

Médigue C, Rouxel T, Vigier P, Hénaut A, Danchin A (1991) Evidence for horizontal gene transfer in *Escherichia coli* speciation. J Mol Biol 222:851–856

Mizuuchi K, Craigie R (1986) Mechanism of bacteriophage Mu transposition. Annu Rev Genet 20:385–429

Naas T, Blot M, Fitch WM, Arber W (1994) Insertion sequence-related genetic variation in resting *Escherichia coli* K-12. Genetics 136: 721–730

Olasz F, Stalder R, Arber W (1993) Formation of the tandem repeat $(IS30)_2$ and its role in IS*30*-mediated transpositional DNA rearrangements. Mol Gen Genet 239:177–187

Sandmeier H (1994) Acquisition and rearrangement of sequence motifs in the evolution of bacteriophage tail fibres. Mol Microbiol 12:343–350

Sekine Y, Ohtsubo E (1989) Frameshifting is required for production of the transposase encoded by insertion sequence *1.* Proc Natl Acad Sci USA 86:4609–4613

Sengstag C, Arber W (1983) IS2 insertion is a major cause of spontaneous mutagenesis of the bacteriophage P1: non-random distribution of target sites. EMBO J 2:67–71

Sengstag C, Arber W (1987) A cloned DNA fragment from bacteriophage P1 enhances IS2 insertion. Mol Gen Genet 206:344–351

Simons RW, Kleckner N (1983) Translational control of IS*10* transposition. Cell 34:683–691

Studier FW, Bandyopadhyay PK (1988) Model for how type I restriction enzymes select cleavage sites in DNA. Proc Natl Acad Sci USA 85:4677–4681

Zerbib D, Polard P, Esconbas J, Galas D, Chandler M (1990) Regulatory role of the IS*1*-encoded InsA protein in transposition. Mol Microbiol 4:471–477