The EMBO Journal vol.15 no.18 pp.5104–5111, 1996

Linear diffusion of the restriction endonuclease *Eco*RV on DNA is essential for the *in vivo* function of the enzyme

Albert Jeltsch, Christian Wenz, Frank Stahl and Alfred Pingoud[1]

Institut für Biochemie, FB Biologie, Justus-Liebig Universität, Heinrich-Buff-Ring 58, 35392 Giessen, Germany

[1]Corresponding author

Linear diffusion along DNA is a mechanism of enhancing the association rates of proteins to their specific recognition sites on DNA. It has been demonstrated for several proteins *in vitro*, but to date in no case *in vivo*. Here we show that the restriction endonuclease *Eco*RV slides along the DNA, scanning ~1000 bp in one binding event. This process is critically dependent on contacts between amino acid residues of the protein and the backbone of the DNA. The disruption of single hydrogen bonds and, in particular, the alteration of electrostatic interactions between amino acid side chains of the protein and phosphate groups of the DNA interfere with or abolish effective sliding. The efficiency of linear diffusion is dependent on salt concentration, having a maximum at 50 mM NaCl. These results suggest that a nonspecific and mobile binding mode capable of linear diffusion is dependent on a subtle balance of forces governing the interaction of the enzyme and the DNA. A strong correlation between the ability of *Eco*RV mutants to slide along the DNA *in vitro* and to protect *Escherichia coli* cells from phage infection demonstrates that linear diffusion occurs *in vivo* and is essential for effective phage restriction.

Keywords: facilitated diffusion/nonspecific DNA binding mode/protein–nucleic acid interaction/restriction modification system/site-directed mutagenesis

Introduction

Type II restriction modification (RM) systems serve to protect bacterial cells from bacteriophage infection (for reviews see Heitman, 1993; Roberts and Halford, 1993). They are composed of two proteins, a restriction endonuclease that cleaves DNA with a very high specificity at defined palindromic sites 4–8 bp in length, and a DNA methyltransferase that modifies these sites within the host DNA and thereby protects the host DNA from cleavage. Depending on the environmental conditions, such systems reduce the probability of bacteriophage infection by two to six orders of magnitude (Korona and Levin, 1993). RM systems are a well understood example of where biological function is predominantly kinetically controlled: if incoming phage DNA is cleaved at the specific sites by the endonuclease before it becomes methylated, infection will be prevented; if, however, methylation takes place prior to cleavage, the phage will infect the cell. Therefore, fast target site location is expected to be one of the most important properties of restriction endonucleases. Interestingly, all restriction enzymes studied so far bind not only to their specific DNA sites but also nonspecifically to DNA (e.g. *Cfr*9I, Siksnys and Pleckaityte, 1993; *Eco*RI, Goppelt *et al.*, 1980; Woodhead and Malcolm, 1980; Lesser *et al.*, 1990; Thielking *et al.*, 1990; *Eco*RV, Taylor *et al.*, 1991; Alves *et al.*, 1995; *Hin*fI, Frankel *et al.*, 1985; *Mbo*II, Sektas *et al.*, 1995; *Rsr*I, Aiken *et al.*, 1991; *Taq*I, Zebala *et al.*, 1992; *Xma*I, Withers and Dunbar, 1995). This property allows in principle for a facilitated target site location, because after nonspecific association anywhere on the DNA the enzyme can stay bound to the DNA and scan the DNA for its recognition site by linear diffusion (Richter and Eigen, 1974; Berg and von Hippel, 1985; von Hippel and Berg, 1989). Although the search for the recognition site by linear diffusion is still a diffusional random process, reduction in dimensionality significantly speeds up target site location compared with a 3-D search in solution (Adam and Delbrück, 1968; Richter and Eigen, 1974). It has been demonstrated by DNA cleavage experiments *in vitro* that the restriction enzyme *Eco*RI makes use of this mechanism (Jack *et al.*, 1982; Ehbrecht *et al.*, 1985; Terry *et al.*, 1985) and that during linear diffusion *Eco*RI follows the helical path of the DNA, pauses at sites that resemble its recognition site and is blocked by irregular DNA structures or ligands bound to the DNA (Jeltsch *et al.*, 1994). The structure of a specific *Eco*RI–DNA complex illustrates the structural basis for linear diffusion because the enzyme enwraps the DNA helix (Kim *et al.*, 1990). As this is observed for all specific complexes between restriction enzymes and DNA known so far (Figure 1), as well as the nonspecific *Eco*RV–DNA complex (Winkler *et al.*, 1993) and many other protein–DNA complexes, e.g. procaryotic and eukaryotic DNA polymerases (Kong *et al.*, 1992; Krishna *et al.*, 1994), linear diffusion is likely to be a widespread phenomenon. For *Escherichia coli* RNA polymerase which also enwraps the DNA (Polyakov *et al.*, 1995), sliding along the DNA has been demonstrated by direct visualization (Kabata *et al.*, 1994).

This study focuses on the restriction enzyme *Eco*RV which recognizes and cleaves GAT↓ATC sites with high specificity. We have quantitatively characterized the efficiency of linear diffusion of *Eco*RV on DNA as well as its dependence on the salt concentration of the cleavage buffer. A special emphasis of this work has been to clarify whether linear diffusion is also important for the *in vivo* function of restriction enzymes. To this end, we have compared the ability of several well characterized *Eco*RV mutants for linear diffusion *in vitro* and the protection of *E.coli* cells from bacteriophage infection by these mutants *in vivo*. Our results demonstrate that *Eco*RV makes use of linear diffusion *in vivo* to effectively protect *E.coli* cells against bacteriophage infection.

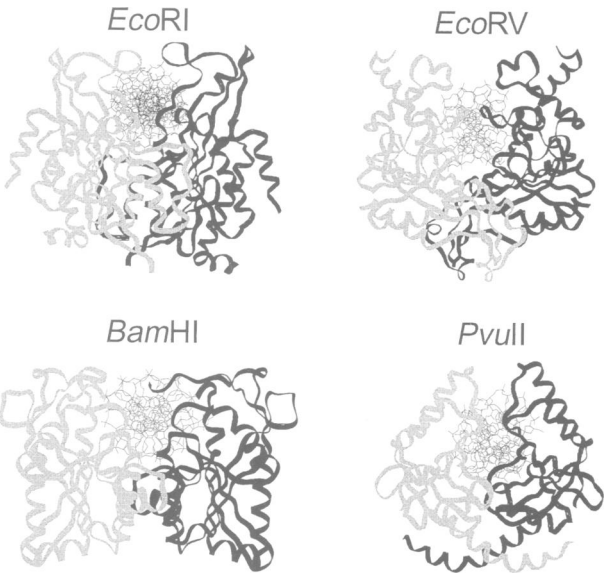

Fig. 1. Structures of four restriction endonucleases in complex with cognate DNA (*Eco*RI, Kim *et al.*, 1990; *Eco*RV, Winkler *et al.*, 1993; Kostrewa and Winkler, 1995; *Pvu*II, Cheng *et al.*, 1994; *Bam*HI, Newman *et al.*, 1995). Note that all proteins enwrap their DNA substrate to give an architecture that is highly suggestive of sliding along the DNA.

Results

It has been the aim of this study to investigate the linear diffusion of *Eco*RV along DNA *in vitro* and *in vivo*. For the *in vitro* experiments, a competitive DNA cleavage assay was employed that is very well suited for this purpose because it is accurate and allows the analysis of many samples in parallel (Jeltsch *et al.*, 1993a, 1994). The assay is based on the competitive cleavage of two DNA substrates and a quantitative analysis of the kinetics of cleavage of both substrates. One substrate is a short oligodeoxynucleotide containing one *Eco*RV site, whereas the other is a macromolecular PCR fragment. Long substrates of different length are generated by PCR using the short oligodeoxynucleotide as one PCR primer. Thus a single *Eco*RV site is introduced into the PCR fragments. Hence, the *Eco*RV site is located in an identical sequence context in all substrates, and the ratio of the cleavage rate constants, $k_{\text{long DNA}}/k_{\text{26mer}}$, is a measure of linear diffusion on the long DNA fragments. As k_{26mer} characterizes the rate of cleavage resulting from direct binding to the sequence of the 26mer, $k_{\text{long DNA}} - k_{\text{26mer}}$ gives the enhancement of the rate of DNA cleavage caused by linear diffusion. Hence, $(k_{\text{long DNA}}/k_{\text{26mer}}) - 1$ (which is equal to $k_{\text{long DNA}} - k_{\text{26mer}}$) is a direct quantitative measure of the efficiency of linear diffusion.

Linear diffusion of EcoRV

Several proteins specifically interacting with DNA have been shown *in vitro* to be able to associate very fast to their recognition site by linear diffusion, including the restriction endonucleases *Eco*RI (Jack *et al.*, 1982; Ehbrecht *et al.*, 1985; Terry *et al.*, 1985; Jeltsch *et al.*, 1994), *Bam*HI (Ehbrecht *et al.*, 1985; Nardone *et al.*, 1987), *Hin*dIII (Ehbrecht *et al.*, 1985), *Eco*RV (Taylor *et al.*, 1991) and *Bss*HI (Berkhout and van Wamel, 1996). To determine, in quantitative terms, the efficiency of linear

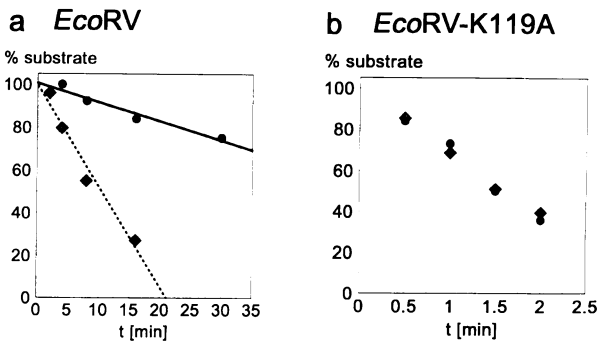

Fig. 2. Time course of the competitive cleavage of the 958mer (◆, ·····) and 26mer (●, —) substrates by 8 nM wild-type *Eco*RV (**a**) or 40 nM K119A mutant (**b**).

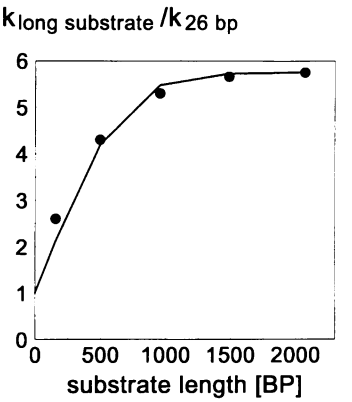

Fig. 3. Length dependence of the efficiency of linear diffusion by *Eco*RV. The data points indicated are the result of at least three independent experiments. All standard deviations are $<\pm 0.2$. The line is the best fit of the data points to Equation 1, yielding a P value of 500 000.

diffusion of *Eco*RV, five DNA substrates of different length (159, 498, 958, 1488 and 2072 bp) were prepared by PCR using the same template, one common upper primer that introduces a single *Eco*RV site in the products and five different lower primers. Therefore the substrates differ in length only, not in the sequences directly flanking the recognition sequence. The substrates were purified and analyzed with respect to their cleavage rates in competition with the 26mer reference substrate, which contains the upper primer as one strand of the duplex. Figure 2a shows the time course of the cleavage of a 958 bp substrate and the reference substrate by *Eco*RV in competition. The 958mer is cleaved 5.6 times faster than the 26mer. The results obtained with all substrates are shown in Figure 3. The rate of DNA cleavage turned out to be strongly dependent on the length of the DNA up to ~1000 bp; beyond that virtually no rate enhancement was observed. This result demonstrates that linear diffusion speeds up the rate of DNA cleavage by *Eco*RV under these conditions.

The contribution of linear diffusion to the rate of association between *Eco*RV and its recognition site on a DNA with a length of L bp can be described by Equation 1 (Ehbrecht *et al.*, 1985):

$$k_{\text{on}}/k_{+1} = \sum_{n=1}^{L} \exp[-(n-s)^2/P]. \qquad (1)$$

Assuming that association can occur with a rate k_{+1} at

Table I. Linear diffusion of wild-type *Eco*RV restriction endonuclease under different NaCl concentrations

Concentration of NaCl (mM)	k_{958mer}/k_{26mer}[a]
0	1.2
20	1.7
50	5.6
100	2.5

[a]All experiments were carried out at least in duplicate. The values given are accurate to within ±0.5.

any site *n* and that the specific binding site is located at position *s*, the overall association rate to the specific site (k_{on}) depends only on the probability *P* that the enzyme diffuses one step along the DNA rather than dissociates from the DNA. This probability is given by the ratio of rate constants for linear diffusion (k_{dif}) and dissociation (k_{off}). *P* corresponds to the average number of linear diffusion steps of the enzyme during one binding event. End effects are ignored in Equation 1 because we do not know at present whether *Eco*RV molecules will be 'reflected' or fall off at the end of a linear DNA molecule, as shown recently for the *Eco*RI DNA methyltransferase (Surby and Reich, 1996). Nevertheless, the equation sufficiently describes the observed length dependence measured for *Eco*RV, which is shown in Figure 3. The data are best fitted by a *P* value of 5×10^5. As linear diffusion is considered to be a thermally driven random walk, the effective average travel length is given by the square root of the number of steps, meaning that *Eco*RV scans on average ~710 bp per binding event under these conditions.

Dependence of linear diffusion on the salt concentration of the cleavage buffer

We have determined the efficiency of linear diffusion under different buffer conditions using the 958 bp substrate as an example. The concentration of NaCl in the cleavage buffer was varied, between 0 and 100 mM at 1 mM $MgCl_2$. The results are shown in Table I. The efficiency of linear diffusion turned out to be strongly dependent on the concentration of NaCl in a complex manner, showing a maximum at 50 mM NaCl. In addition, we have determined linear diffusion with the 958mer at high NaCl concentrations (100 mM) under more physiological Mg^{2+} concentrations (5 mM), and found that under these conditions linear diffusion also speeds up the rate of DNA cleavage by *Eco*RV (k_{958mer}/k_{26mer} = 4.5).

Linear diffusion of EcoRV variants

To examine the importance of the integrity of the protein–DNA interface on the ability of *Eco*RV to slide along the DNA, we have analyzed the ratio of cleavage rates (k_{958mer}/k_{26mer}) of various *Eco*RV mutants. For experimental reasons, only mutants could be used whose DNA cleavage rates are not reduced by more than two orders of magnitude with respect to the wild-type enzyme, because with less active enzymes very high protein concentrations would be necessary to carry out the cleavage experiments, which would obscure the effect of linear diffusion. We have selected *Eco*RV mutants in which contacts to the phosphodiester backbone of the DNA are impaired (S41A, T93A, T94A, Y95F, T111A, S112A, K119A, R140A, R221A, S223A and R226A), the rationale being that backbone contacts are expected to be of particular importance for linear diffusion. These mutants have been characterized previously and their properties have been described (Wenz *et al.*, 1996). We also analyzed an *Eco*RV variant in which one catalytically important aspartic acid residue (Asp90) in only one subunit of the dimeric enzyme is replaced by alanine (wt/D90A). Details concerning the purification and properties of this enzyme variant are described elsewhere (Stahl *et al.*, 1996). In this mutant the electrostatic potential of *Eco*RV is minimally disturbed because only one charge is deleted. The results of these competitive cleavage experiments with the *Eco*RV variants, one of which is shown in Figure 2b, are compiled in Table II. Many mutants show a reduced efficiency of linear diffusion; some of them are virtually unable to diffuse along the DNA, e.g. the K119A and wt/D90A variants. One mutant, however, diffuses better than wild-type *Eco*RV (R221A).

Table II. Linear diffusion and *in vivo* activity of wild-type *Eco*RV restriction endonuclease and various mutants

Variant	k_{958mer}/k_{26mer}[a]	λ phage titer (p.f.u./μl)[b]
Wild type	5.6	6.2×10^2
S41A	4.8	1.7×10^3
T93A	1.8	5.1×10^5
T94A	2.6	1.8×10^4
Y95F	4.0	2.2×10^3
T111A	3.3	3.2×10^4
S112A	4.8	3.1×10^2
K119A	1.0	6.2×10^5
R140A	1.7	3.1×10^5
R221A	8.7	4.9×10^2
S223A	1.9	2.3×10^4
R226A	1.5	5.3×10^5
wt/D90A	1.1	ND

ND, not determinable.
[a]All experiments were carried out at least in triplicate. The values given are accurate within at least ±0.2.
[b]All determinations were carried out at least in triplicate. The values given are accurate within at least a factor of five. The phage titer determined with *E.coli* C600 control cells only containing the methyltransferase plasmid was 8.2×10^5.

In vivo measurement of EcoRV activity under physiological conditions

Linear diffusion along the DNA has been demonstrated to speed up the DNA association rates of many proteins interacting with DNA *in vitro*, including many restriction endonucleases. However, the importance of this property for the function of the restriction enzymes *in vivo* remained unclear (cf. Discussion). As it appears nearly impossible to simulate all the relevant *in vivo* parameters *in vitro* (the nature and concentration of monovalent and bivalent cations and anions, as well as polyamines, the effective concentration of nuclease, the influence of other DNA-binding proteins, molecular crowding, etc.), we decided to set up an assay based on the *in vivo* function of *Eco*RV to protect *E.coli* cells against phage infection. For this purpose, *E.coli* C600 cells were transformed with two compatible plasmids, the pLMB4422 plasmid that encodes the *Eco*RV methyltransferase and one of the plasmids of the pHisRV series coding for different *Eco*RV mutants. In these cells the expression levels of all endonuclease

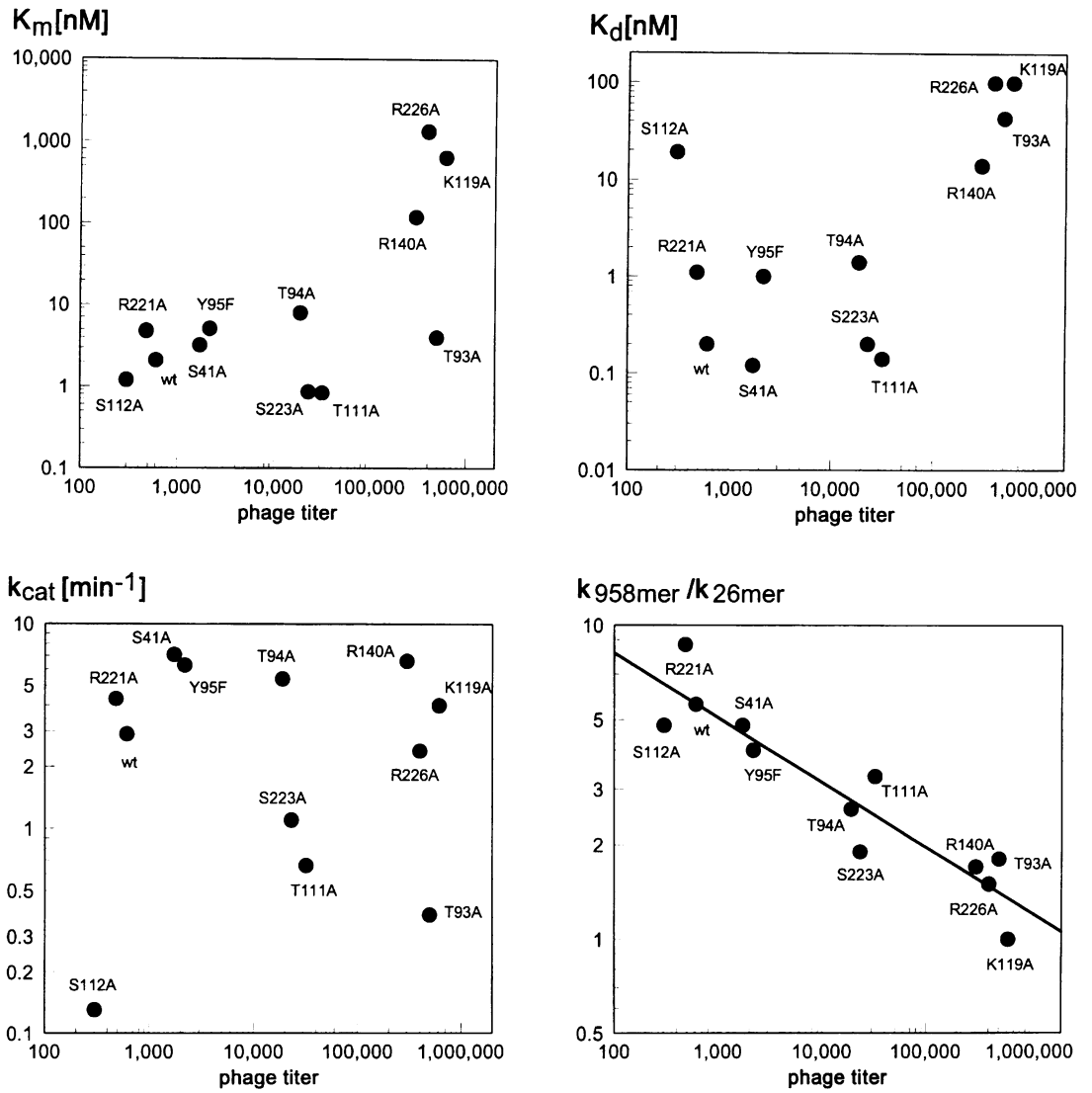

Fig. 4. Correlation of the protection of *E.coli* cells against phage infection by wild-type *Eco*RV and various *Eco*RV mutants with *in vitro* properties of the variants. The K_m, K_d and k_{cat} values are taken from Wenz *et al.* (1996); k_{958mer}/k_{26mer} and phage titers are from Table II.

mutants were similar, with the total concentration being in the order of 1000 molecules per cell (data not shown). Subsequently, the cells that now contain functional (mutant) *Eco*RV restriction modification systems were infected with λ-phages. The phage titers shown in Table II demonstrate that wild-type *Eco*RV efficiently protects the cells from infection by approximately three orders of magnitude, compared with *E.coli* cells not expressing *Eco*RV. This result means that only every 1000th phage invading an *E.coli* cell escapes from nucleolytic cleavage and becomes methylated. As all its descendants are also methylated, surrounding cells are no longer protected by the *Eco*RV endonuclease and a plaque results. It turned out that some mutants had a very similar protection factor as wild-type *Eco*RV, whereas others did not have any protective effect. To find out which of the enzymatic properties of the mutants are relevant for restriction *in vivo*, we analyzed to what extent the k_{cat}, K_m, K_d (taken from Wenz *et al.*, 1996) or k_{958mer}/k_{26mer} values (Table II) were correlated to the phage protection data. The steady state kinetic parameters of the *Eco*RV variants (K_m and k_{cat}) had been determined using a 20mer oligodeoxynucleotide substrate under conditions similar to those employed here for the plasmid DNA cleavage experiments. Gelshift experiments using a 382mer DNA substrate in the presence of Ca^{2+} (Vipond and Halford, 1995) had been employed to determine specific DNA-binding constants of the mutant proteins (K_d). The plots of K_m versus phage titer, k_{cat} versus phage titer and K_d versus phage titer do not show clear correlations (Figure 4). However, a very good correlation is obtained with the k_{958mer}/k_{26mer} values. In this case a correlation coefficient (R) of –0.92 is observed. Monte Carlo simulations demonstrated that the correlation observed between linear diffusion and phage restriction has a probability of only 3×10^{-5} (0.003%) of occurring by chance. We conclude from this finding that linear diffusion takes place *in vivo* and is essential for the *in vivo* function of *Eco*RV.

Discussion

Linear diffusion of proteins on DNA has been demonstrated for many proteins *in vitro*, including repressors (Winter and von Hippel, 1981; Winter *et al.*, 1981),

polymerases (Singer and Wu, 1987), DNA-methyltransferases (Nardone *et al.*, 1987; Surby and Reich, 1996), repair endonucleases (Lloyd *et al.*, 1980; Hamilton and Lloyd, 1989), restriction endonucleases (Jack *et al.*, 1982; Ehbrecht *et al.*, 1985; Terry *et al.*, 1985) and exonucleases (Lu and Gray, 1995). It has been seen to be an effective means for fast target site location, which could be of importance *in vivo*. The most detailed analyses of the process of linear diffusion have been carried out for restriction endonucleases, in particular *Eco*RI (Jack *et al.*, 1982; Ehbrecht *et al.*, 1985; Terry *et al.*, 1985; Jeltsch *et al.*, 1994), for which it was shown to be a random diffusional walk on the DNA in which the enzyme follows the helical pitch of the DNA. The sliding movement is slowed down when sequences are encountered that resemble the recognition sequence, or interrupted when other proteins, drugs, etc., are bound to the DNA. Two kinds of question usually arise in discussions on linear diffusion, namely (i) What kind of interactions between protein and DNA are responsible to enable the protein to slide along the DNA?, and, more critical, (ii) Is linear diffusion only an *in vitro* phenomenon that does not occur *in vivo*? We have attempted to address both questions here.

Linear diffusion of EcoRV and EcoRV variants in vitro

The *in vitro* activity of *Eco*RV is strongly dependent on linear diffusion. In a comparison of the cleavage rates of a 26mer reference substrate and a 958 bp PCR fragment, the long substrate was cleaved up to 5.6 times faster, depending on the cleavage conditions. The cleavage rates measured for five substrates of different length demonstrate that *Eco*RV scans on average almost 1000 bp before it dissociates from the DNA. This value is comparable with results obtained with *Eco*RI (Ehbrecht *et al.*, 1985). Under these conditions ~60% of all *Eco*RV molecules reach their recognition sequence after an association anywhere to the 958mer substrate. As this substrate is cleaved 5.6 times faster than the 26mer, one can estimate that the k_{958mer}/k_{26mer} would be ~9.3 if all *Eco*RV molecules reached the specific site after a nonspecific association. This value is smaller than one might have been expected because the 958mer is 37 times longer than the 26mer. However, one has to keep in mind that linear diffusion of a protein along DNA is much slower than diffusion in solution. For example, it has been estimated for *Eco*RI that linear diffusion is at least two orders of magnitude slower than 3-D diffusion (Ehbrecht *et al.*, 1985). Nevertheless, the reduction in dimensionality speeds up target site location.

We have investigated the nature of the nonspecific binding mode of *Eco*RV on DNA supporting linear diffusion by analyzing several enzyme variants. The results demonstrate that electrostatic interactions are very important for linear diffusion because in those four mutants having the lowest k_{958mer}/k_{26mer} value charged residues are always replaced (K119A, wt/D90A, R226A and R140A). In three of these variants positively charged residues are replaced. The finding that these mutants are impaired in linear diffusion is in agreement with the model that nonspecific binding and linear diffusion of *Eco*RV is strongly dependent on electrostatic attraction. Moreover, the efficiency of linear diffusion is reduced at salt concentrations >50 mM NaCl, again suggesting that electrostatic interactions are important for this process. In agreement with these results, the electrostatic potential of *Eco*RV resembles an electrostatic trap that is well suited to accommodate DNA (Figure 5).

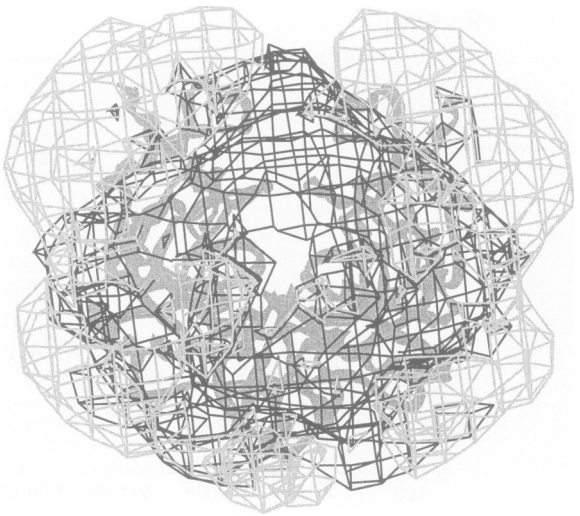

Fig. 5. Electrostatic potential map of *Eco*RV. The isopotential lines of +3 (black) and –3 kcal/mol (gray) are shown. The map was calculated with Sybyl 6.0 (Tripos, San Diego, CA) on the basis of the specific *Eco*RV DNA complex (Winkler *et al.*, 1993) but omitting the DNA. The orientation of *Eco*RV is the same as in Figure 1. The nonspecific *Eco*RV–DNA complex (Winkler *et al.*, 1993) has a similar general architecture to the specific one, although the DNA is contacted less tightly. As some loops of the protein in this complex are not ordered, it cannot be used for a meaningful calculation of electrostatic potential.

The results obtained with mutants in which single hydrogen bonds between the enzyme and the phosphodiester backbone of the DNA are deleted show that these variants also have a reduced ability to diffuse along the DNA. Considering, for example, the S223A mutant, the removal of only one hydroxyl group in each subunit of *Eco*RV which contacts a phosphate group of the DNA causes a decrease in linear diffusion by >80% [compare $(k_{958mer}/k_{26mer})-1$ for wild type (4.6) with S223A mutant (0.9)]. This result demonstrates that hydrogen bonds also contribute to the nonspecific binding mode capable of linear diffusion.

However, not all results obtained are in agreement with such a 'strong binding-efficient linear diffusion' model.

(i) At 1 mM $MgCl_2$ linear diffusion is maximal at a concentration of 50 mM NaCl. This result appears surprising as it has been shown for many DNA-binding proteins that lowering the salt concentration increases the nonspecific DNA-binding constant because of a decrease in k_{off}. This should directly increase the efficiency of linear diffusion.

(ii) The R221A mutant is superior to the wild-type enzyme regarding linear diffusion, although positive charges are deleted in the mutant.

(iii) In the heterodimer wt/D90A, taking away one negative charge nearly completely abolishes linear diffusion. This amino acid residue is implicated in binding of the Mg^{2+} cofactor in the specific *Eco*RV–DNA complex (Selent *et al.*, 1992; Jeltsch *et al.*, 1993b; Kostrewa and Winkler, 1995). However, presumably, this Mg^{2+}-binding site is not occupied in nonspecific complexes (Vermote

and Halford, 1992; Vipond and Halford, 1993) and, therefore, the negative charge of Asp90 is most probably not compensated in nonspecific complexes. Then one would expect that Asp90 repels the DNA and decreases the binding constant, an assumption that is supported by the finding that the homodimeric D90A mutant, in which both subunits are mutated, binds DNA more strongly than wild-type *Eco*RV (Thielking *et al.*, 1992).

To explain these results we propose the following model. The enzyme is considered to be a multidentate ligand traveling along a linear lattice. Electrostatic attraction serves to hold the enzyme on the DNA, controlling the dwell time via k_{off}. As the surface of the DNA is not smooth but structured, the enzyme can interact with the DNA 'in phase', so that, for example, hydrogen bonds between the enzyme and the phosphate groups can form, and 'out of phase' in an intermediate position with no or fewer hydrogen bonds. This leads to an increase in total free energy which determines the activation energy for moving on by 1 bp, and, hence, k_{dif}. Linear diffusion depends on a tight but mobile binding mode. Consequently, according to Equation 1, the relative values of k_{dif} and k_{off} determine the efficiency of linear diffusion. Hydrogen bonds formed between *Eco*RV and the phosphate groups of the DNA on the one hand increase DNA binding affinity and, consequently, reduce k_{off}. On the other hand, they disturb linear diffusion because they have to be broken before sliding can occur. Hence, such interactions increase the activation barrier of sliding from one base pair to the next, resulting in a decrease in k_{dif}. Therefore, a compromise has to be made between too low an affinity, which would prevent linear diffusion by reducing k_{off}, and binding that is too tight, which also would prevent linear diffusion because of a reduced k_{dif} (one might compare this with 'strong friction'). As the enzyme enwraps the DNA, electrostatic attraction on the one side can be balanced by electrostatic attraction on the other or by electrostatic repulsion on the same side. The contradictory requirements for efficient linear diffusion (tight but mobile binding) can possibly explain the NaCl dependence of linear diffusion of *Eco*RV shown here, which at 1 mM $MgCl_2$ has a maximum at 50 mM NaCl. It is probable that DNA binding by *Eco*RV is too tight at low concentrations of NaCl, whereas it is too weak at high NaCl concentrations. Moreover, the wt/D90A mutant at normal salt concentrations shows a similar behavior to wild-type *Eco*RV at low salt concentrations because in both cases strong DNA binding prevents linear diffusion. In the R221A mutant a slightly reduced DNA binding affinity might allow for more efficient linear diffusion because of an increase in k_{dif}.

Taken together, these results show that the ability of *Eco*RV to slide along the DNA is a highly evolved property of the enzyme based mainly on electrostatic interactions, but to a lower degree also on hydrogen bonds between *Eco*RV and the DNA backbone that can easily be disturbed by very subtle changes in the protein–DNA interface. The effects of electrostatic contacts for the interaction between *Eco*RV and nonspecific DNA are more complicated than can be explained by a simple electrostatic attraction model. It rather appears that an electrostatic trap formed by a balance of attractive and repulsive forces is necessary for the enzymes to be able to bind DNA firmly and at the same time to slide along the DNA. The intricate nature of this mobile binding mode was demonstrated for *Eco*RI by the effect of star sites (i.e. sites that differ from the recognition site by only 1 bp) on linear diffusion. Star sites, which are bound by the enzyme more strongly than average DNA sites, slow down linear diffusion to a degree that is roughly correlated with their binding constant to *Eco*RI (Jeltsch *et al.*, 1994).

In vivo contribution of linear diffusion to the activity of EcoRV

Ever since it was demonstrated that linear diffusion is important for the activity of restriction enzymes *in vitro* (Jack *et al.*, 1982; Ehbrecht *et al.*, 1985; Terry *et al.*, 1985) it has been debated whether this process is of importance *in vivo*. On the one hand molecular crowding in the cytosol makes 3-D diffusion even slower than *in vitro*, raising the need for rate enhancement of target site location. On the other hand, it has been shown for *Eco*RI that linear diffusion is strongly disturbed, if not prevented, by other proteins bound specifically at other sites on the DNA (Jeltsch *et al.*, 1994) or nonspecifically to the target DNA (Ehbrecht *et al.*, 1985). As DNA *in vivo* is usually tightly associated with a variety of different proteins, one has to question whether linear diffusion occurs over longer distances *in vivo*. It must be considered, however, that the biological target of restriction endonucleases is phage DNA that has just invaded the cell, and one would not expect that many proteins are bound to the DNA at this stage. To find out experimentally if linear diffusion is important for a restriction enzyme *in vivo*, we compared the efficiency of linear diffusion and the ability to protect *E.coli* cells from phage infection of wild-type *Eco*RV and 12 *Eco*RV mutants. In these mutants, individual amino acid residues that contact phosphate groups of the DNA are exchanged. These enzyme variants have catalytic activities ranging from wild-type activity to ~1/100 (Wenz *et al.*, 1996). It turned out that no clear correlation exists between K_m, k_{cat} or K_d values of the enzyme variants and phage restriction. In contrast, the efficiency of linear diffusion of the mutants was strongly correlated with the *in vivo* activity of the enzyme variants, supporting the idea that linear diffusion indeed takes place in the cell and is essential for the biological function of *Eco*RV, although our data cannot provide precise information regarding the efficiency of linear diffusion *in vivo*, e.g. whether, like *in vitro*, ~1000 bp are scanned.

The correlation between the efficiency of *Eco*RV mutants to diffuse on the DNA *in vitro* and their ability to protect *E.coli* cells from phage infection can be rationalized by considering the molecular events following the invasion of phage DNA in bacterial cells containing a functional type II RM system. As both enzymes are present (the restriction endonuclease and the methyltransferase), they compete for their substrates. The outcome of the competition is mainly kinetically controlled. If the phage DNA is methylated prior to cleavage, the phage will infect the cell. If the restriction enzyme finds its target sites before the methyltransferase, it can cleave the phage DNA and prevent infection. Thus, very fast target site location is of fundamental importance for the function of restriction enzymes. Other properties are not as critical for the *in vivo* function. For example, it may be unimportant

whether cleavage is fast or slow, as long as the enzyme firmly binds to its target site after a fast association, because the enzyme by itself protects the noncleaved site from methylation. This argument could explain why k_{cat} appears to be of little importance *in vivo*. Nevertheless, a reasonable DNA-binding and DNA cleavage activity is a prerequisite for a restriction enzyme variant to restrict phages. As in this study, we investigated only mutants having at least a 1% wild-type activity; perhaps the DNA cleavage and binding activities of all of these mutants were sufficient enough not to become limiting in the phage restriction test *in vivo*.

In conclusion, we have demonstrated that the linear diffusion of *Eco*RV is a highly evolved property that can easily be disturbed by amino acid substitutions. A clear correlation of the efficiency of linear diffusion of various *Eco*RV mutants with the ability of these mutants to protect *E.coli* cells from phage infection strongly suggests that linear diffusion is very important for the *in vivo* function of the restriction endonuclease. It is unlikely that it will ever be possible to obtain direct proof for linear diffusion of a protein on DNA *in vivo*; one rather has to rely on correlations of one sort or another. The positive correlation between phage restriction efficiency *in vivo* and linear diffusion *in vitro* is presumably the best kind of evidence one could expect. As other restriction enzymes have been demonstrated to be capable of sliding along the DNA *in vitro*, it is probable that they also make use of linear diffusion *in vivo*. It is likely that this is also the case for other proteins which interact with DNA and whose reactions are kinetically controlled. For example, the post-replicative dam mismatch repair system closely resembles type II RM systems studied here: whether a misincorporated nucleotide is repaired or not depends on whether the repair endonuclease locates its target prior to the methylation of nearby dam sites by the dam methyltransferase (for reviews see Messer and Noyer-Weidner, 1988; Modrich, 1991).

Materials and methods

Protein construction and purification

For this study wild-type *Eco*RV and the following *Eco*RV mutants were used: S41A, T93A, T94A, Y95F, T111A, S112A, K119A, R140A, R221A, S223A and R226A. In these mutants, amino acids that form contacts with the phosphate groups of the DNA are exchanged. The construction and properties of the mutants have been described by Wenz *et al.* (1996). All proteins carry an affinity tag of six histidines on their N-terminus to facilitate purification. Protein expression and purification were performed as described previously (Wenz *et al.*, 1994). Furthermore, the T186S mutant was used. This enzyme contains an amino acid exchange in the DNA recognition loop of *Eco*RV and is catalytically almost inactive (Thielking *et al.*, 1991). This variant was also purified as a His_6-tagged protein. In addition, we investigated the linear diffusion of a heterodimeric wt/D90A enzyme. In this *Eco*RV variant a mutation (Asp90→Ala) is introduced in only one subunit of the dimeric protein, producing an artificial heterodimer. The mutation inactivates the catalytic center of the corresponding subunit, whereas the second catalytic center retains its catalytic activity. Purification and properties of this and other artificial heterodimers of *Eco*RV have been described by Stahl *et al.* (1996). Briefly, the heterodimer is produced *in vivo* by coexpression of the different subunits and purified using two affinity chromatography steps directed towards two different affinity tags on both subunits.

Oligodeoxynucleotides

Oligodeoxynucleotides were synthesized on solid support with a Milligen Cyclone DNA synthesizer using β-cyanoethylphosphoramidites obtained from MWG Biotech (Ebersberg, Germany). Biotin-phosphoramidites and 'amino-linker' containing phosphoramidites were obtained from Millipore. Oligodeoxynucleotides containing an 'amino-linker' on their 5′ ends were labeled with fluorescein isothiocyanate (Sigma) or digoxygenin-3-*O*-methylcarbonyl-ε-aminocaproic acid-*N*-hydroxy-succinimide ester (Boehringer-Mannheim), as described previously (Jeltsch *et al.*, 1993a), to produce oligodeoxynucleotides whose 5′ ends were labeled with fluorescein or digoxygenin.

PCR

DNA substrates of different length were used to investigate the efficiency of linear diffusion of *Eco*RV. These substrates were generated by PCR using a pAT153–plasmid template, different biotinylated lower primers and one common upper primer d (amino-linker-TGGCGCC**GATAT-**CGCCGACATCAAAG) labeled on its 'amino-linker' with fluorescein prior to the PCR. The lower primers were chosen to amplify fragments 158, 498, 958, 1488 and 2092 bp in length. In all of these substrates, the *Eco*RV cleavage site is flanked by a short arm comprising 7 bp and a long arm comprising 145–2079 bp. PCR was carried out in a thermocycler varius V45 (Landgraf, Hannover). 1.5 ng template DNA were used in a 100 µl PCR with 0.4 µM of each primer and 0.2 mM dNTPs. The reaction was carried out with 5 U *Taq* polymerase (Promega) in Taq 1× reaction buffer (cycle 1, 180 s at 93°C; cycle 2, 30 s at 93°C, 60 s at 50°C and 60 s at 72°C, cycle 2 was repeated 35 times; cycle 37, 30 s at 93°C, 30 s at 54°C and 200 s at 72°C). All PCRs yielded one homogeneous product. The DNA was purified using a PCR spin prep kit (Qiagen). The concentrations of the products were estimated from ethidium bromide-stained gels.

Generation of a reference substrate

To generate a reference substrate, the 26 nucleotide-long upper PCR primer was labeled with digoxygenin (see above) and annealed to a complementary oligodeoxynucleotide which had been biotinylated on its 5′ end. The resulting 26 bp substrate was used as an internal reference in the competitive cleavage assay.

Competitive cleavage assay of two substrates

Cleavage reactions were carried out with one long DNA fragment and the 26 bp reference substrate in competition by using an ELISA-based assay procedure as described previously (Jeltsch *et al.*, 1993a). Briefly, the assay employed oligodeoxynucleotides labeled with different haptens (long substrates, fluorescein; 26 bp reference substrate, digoxigenin) on one strand. A 5′ biotin label of the complementary strand was used to separate products and uncleaved substrates on avidin-coated microtiter plates by washing. Cleavage of the long substrate and the 26 bp standard was carried out in competition in one tube; the disappearance of both substrates was detected by two separate series of ELISA reactions. Cleavage reactions were carried out in buffer conditions that were demonstrated to promote the linear diffusion of *Eco*RI (Ehbrecht *et al.*, 1985): 1 mM $MgCl_2$, 20 mM Tris–HCl, pH 7.5, 50 µg/ml bovine serum albumin at 21°C. If not otherwise stated, NaCl concentrations were 50 mM. Substrate concentrations were 10 (long substrates) and 20 nM (26mer). Depending on their activity, the concentrations of the *Eco*RV mutants in the assay were 8 (wild type, S41A, Y95F, T111A, S112A, R221A and S223A) or 40 nM (T93A, T94A K119A, R140A and R226A).

Phage restriction assay

For the phage restriction assay, *E.coli* C600 cells harboring the plasmids pLBM4422 (Thielking *et al.*, 1991) and pHisRV (Wenz *et al.*, 1994) were used, which code for the *Eco*RV methyltransferase and for wild-type or mutant *Eco*RV endonucleases, respectively. The methyltransferase and the endonuclease genes are constitutively expressed at a low level. λ_{O125} phages were a kind gift from Dr G.Hobom (Institut für Mikrobiologie und Molekularbiologie, Universität Giessen, Germany). λ phage titers were determined using standard techniques (Ausubel *et al.*, 1995).

Correlation analysis

To find out whether and to what extent correlations between two properties of wild-type and mutant *Eco*RV enzymes (e.g. between phage restriction and linear diffusion) were significant, Monte Carlo simulations were carried out. First, the *R* value of the correlation was calculated, which gives a quantitative estimate of the degree of correlation. Then, the data from both data sets were randomly paired. For example, the phage restriction value of mutant X becomes paired with the linear diffusion factor measured for mutant Y. The linear diffusion factor measured for mutant X becomes assigned randomly to the phage

restriction value measured for another variant, and so on. Then the *R*-factor of the random data pairs was calculated. This shuffling was repeated 100 000 times, and the number of *R* values equal to or higher than the original *R* value was counted.

Acknowledgements

We thank Prof. Dr J.Alves for stimulating discussions. Technical assistance by Ms H.Büngen is gratefully acknowledged. This work was supported by the Deutsche Forschungsgemeinschft (Pi 122/12-1) and the Fonds der Chemischen Industrie.

References

Adam,G. and Delbrück,M. (1968) In Rich,A. and Davidson,N. (eds), *Structural Chemistry and Molecular Biology*, Freeman and Co., San Fransisco, CA, pp. 198–215.

Aiken,C.R., Fisher,E.W. and Gumport,R.I. (1991) *J. Biol. Chem.*, **266**, 19063–19069.

Alves,J., Selent,U. and Wolfes,H. (1995) *Biochemistry*, **34**, 11191–11197.

Ausubel,F.M. *et al.* (1995) *Current Protocols in Molecular Biology*. John Wiley and Sons, New York, Vol. 1.

Berg,O.G. and von Hippel,P.H. (1985) *Annu. Rev. Biophys. Biophys. Chem.*, **14**, 131–160.

Berkhout,B. and van Wamel,J. (1996) *J. Biol. Chem.*, **271**, 1837–1840.

Cheng,X., Balendiran,K., Schildkraut,I. and Anderson,J.E. (1994) *EMBO J.*, **13**, 3927–3935.

Ehbrecht,H.-J., Pingoud,A., Urbanke,C., Maass,G. and Gualerzi,C. (1985) *J. Biol. Chem.*, **260**, 6160–6166.

Frankel,A.D., Ackers,G.K. and Smith,H.O. (1985) *Biochemistry*, **24**, 3049–3054.

Goppelt,M., Pingoud,A., Maass,G., Mayer,H., Koster,H. and Frank,R. (1980) *Eur. J. Biochem.*, **104**, 101–107.

Hamilton,R.W. and Lloyd,R.S. (1989) *J. Biol. Chem.*, **264**, 17422–17427.

Heitman,J. (1993) In Setlow,J.K. (ed.), *Genetic Engineering*. Plenum Press, New York, pp. 57–108.

Jack,W.E., Terry,B.J. and Modrich,P. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 4010–4014.

Jeltsch,A., Fritz,A., Alves,J., Wolfes,H. and Pingoud,A. (1993a) *Anal. Biochem.*, **213**, 234–240.

Jeltsch,A., Alves,J., Wolfes,H., Maass,G. and Pingoud,A. (1993b) *Proc. Natl Acad. Sci. USA*, **90**, 8499–8503.

Jeltsch,A., Alves,J., Wolfes,H., Maass,G. and Pingoud,A. (1994) *Biochemistry*, **33**, 10215–10219.

Kabata,H., Kurosawa,O., Arai,I., Washizu,M., Margarson,S.A., Glass,R.E. and Shimamoto,N. (1994) *Science*, **262**, 1561–1563.

Kim,Y., Grable,J.C., Love,R., Greene,P.J. and Rosenberg,J.M. (1990) *Science*, **249**, 1307–1309.

Kong,X.P., Onrust,R., O'Donnell,M. and Kuriyan,J. (1992) *Cell*, **69**, 425–437.

Korona,R. and Levin,B.R. (1993) *Evolution*, **47**, 556–575.

Kostrewa,D. and Winkler,F.K. (1995) *Biochemistry*, **34**, 683–696.

Krishna,T.S., Kong,X.P., Gary,S., Burgers,P.M. and Kuriyan,J. (1994) *Cell*, **79**, 1233–1243.

Lesser,D.R., Kurpiewski,M.R. and Jen-Jacobson,L. (1990) *Science*, **250**, 776–786.

Lloyd,R.S., Hanawalt,P.C. and Dodson,M.L. (1980) *Nucleic Acids Res.*, **8**, 5113–5127.

Lu,T. and Gray,H.B.,Jr (1995) *Biochem. Biophys. Acta*, **1251**, 125–128.

Messer,W. and Noyer-Weidner,M. (1988) *Cell*, **54**, 735–737.

Modrich,P. (1991) *Annu. Rev. Genet.*, **25**, 229–253.

Nardone,G., George,J. and Chirikjian,J.G. (1987) *J. Biol. Chem.*, **261**, 12128–12133.

Newman,M., Strzelecka,T., Dorner,L.F., Schildkraut,I. and Aggarwal, A.K. (1995) *Science*, **269**, 656–663.

Polyakov,A., Severinova,E. and Darst,S.A. (1995) *Cell*, **83**, 365–373.

Richter,P.H. and Eigen,M. (1974) *Biophys. Chem.*, **2**, 255–263.

Roberts,R.J. and Halford,S.E. (1993) In Linn,S.M., Lloyd,R.S. and Roberts,R.J. (eds), *Nucleases*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 35–88.

Sektas,M., Kaczorowski,T. and Podhajska,A.J. (1995) *Gene*, **157**, 181–185.

Selent,U. *et al.* (1992) *Biochemistry*, **31**, 4808–4815.

Singer,P. and Wu,C.W. (1987) *J. Biol. Chem.*, **262**, 14178–14189.

Siksnys,V. and Pleckaityte,M. (1993) *Eur. J. Biochem.*, **217**, 411–419.

Stahl,F., Wende,W., Jeltsch,A. and Pingoud,A. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 1781–1795.

Surby,M.A. and Reich,N.O. (1996) *Biochemistry*, **35**, 2201–2208.

Taylor,J.D., Badcoe,I.G., Clarke,A.R. and Halford,S.E. (1991) *Biochemistry*, **30**, 8743–8753.

Terry,B.J., Jack,W.E. and Modrich,P. (1985) *J. Biol. Chem.*, **260**, 13130–13137.

Thielking,V., Alves,J., Fliess,A., Maass,G. and Pingoud,A. (1990) *Biochemistry*, **29**, 4682–4691.

Thielking,V., Selent,U., Köhler,E., Wolfes,H., Pieper,U., Geiger,R., Urbanke,C., Winkler,F.K. and Pingoud,A. (1991) *Biochemistry*, **30**, 6416–6422.

Thielking,V., Selent,U., Köhler,E., Landgraf,Z., Wolfes,H., Alves,J. and Pingoud,A. (1992) *Biochemistry*, **31**, 3727–3732.

Vermote,C.L.M. and Halford,S.E. (1992) *Biochemistry*, **31**, 6082–6089.

Vipond,B. and Halford,S.E. (1993) *Mol. Microbiol.*, **9**, 225–231.

Vipond,I.B. and Halford,S.E. (1995) *Biochemistry*, **34**, 1113–1119.

von Hippel,P.H. and Berg,O.G. (1989) *J. Biol. Chem.*, **264**, 675–678.

Wenz,C., Selent,U., Wende,W., Jeltsch,A., Wolfes,H. and Pingoud,A. (1994) *Biochim. Biophys. Acta*, **1219**, 73–80.

Wenz,C., Jeltsch,A. and Pingoud,A. (1996) *J. Biol. Chem.*, **271**, 5565–5573.

Winkler,F.K. *et al.* (1993) *EMBO J.*, **12**, 1781–1795.

Winter,R.B. and von Hippel,P.H. (1981) *Biochemistry*, **20**, 6961–6977.

Winter,R.B., Berg,O.G. and von Hippel,P.H. (1981) *Biochemistry*, **20**, 6948–6960.

Withers,B.B. and Dunbar,J.C. (1995) *Nucleic Acids Res.*, **23**, 3571–3577.

Woodhead,J.L. and Malcolm,A.D.B. (1980) *Nucleic Acids Res.*, **8**, 389–402.

Zebala,J., Choi,J. and Barany,F. (1992) *J. Biol. Chem.*, **267**, 8097–8105.

Received on April 24, 1996; revised on June 7, 1996