Sequence-specific responses of restriction endonucleases to bromodeoxyuridine substitution in mammalian DNA

John Petruska and Diane Horn

Molecular Biology Division, University of Southern California, Los Angeles, CA 90089, USA

Received 8 September 1982; Revised 9 March 1983; Accepted 18 March 1983

ABSTRACT

Substitution of BrdU for dT in mammalian DNA alters the rates of DNA cleavage by restriction endonucleases in a manner that can be related to the specificity of cleavage. A formula is proposed that describes inhibitory and stimulatory contributions arising from the substitution of a Br atom for the CH₃ group on T. The larger Br atom is postulated to sterically hinder the nuclease from binding to adjacent groups in the DNA cleavage site, while allowing a tighter binding to itself. The inhibition caused by steric hindrance is predicted to vary inversely with distance from the point of cleavage, whereas the stimulation caused by tighter binding is predicted to be independent of distance. The resultant formula gives a good fit to the data obtained for thirteen different restriction nucleases of known specificity. The parameters in the formula appear to be simple functions of ionic strength. This formula can be used to predict the effect of BrdU substitution on any endonuclease whose specificity of cleavage is known.

INTRODUCTION

Total replacement of thymidine (dT) in DNA by 5-bromodeoxyuridine (BrdU) can be obtained in the Syrian hamster melanoma cell line HAB (1). This results in the substitution of a Br atom for the CH₃ group on each T in the wide groove of the DNA double helix. The resultant DNA provides a means of testing the sensitivity of proteins that recognize specific base sequences in the helix. Proteins of known sequence specificity, in particular lac repressor (2) and restriction endonucleases (3-5), show distinct changes in binding or cleavage rate. The responses of these highly specific proteins are indicative of the effects of BrdU at the DNA level in cells. Understanding the basis of the responses could be relevant to an understanding of the effects of BrdU *in vivo*, e.g., inhibition of cell differentiation, induction of viruses and the inhibition or stimulation of specific gene products (for review, see reference 6).

The restriction endonucleases provide many examples of sequence-specific cleavage of DNA (7). By comparing the rates of cleavage in DNA with and with-

out BrdU, it is possible to measure the effects of BrdU substitution in the cleavage site of each enzyme. Previous studies have shown that a number of enzymes are inhibited to different extents by BrdU (3-5), whereas one enzyme is stimulated. To explain these differences, we have suggested that the ability of BrdU to inhibit or stimulate cleavage varies systematically with position in the cleavage site (5). In order to test this hypothesis, we have examined a total of thirteen different restriction endonucleases with the same preparations of 100% BrdU substituted (B) and unsubstituted (T) DNA. The results have led us to formulate a general equation for the inhibitory and stimulatory contributions as a function of the number and placement of T's in the cleavage site.

MATERIALS AND METHODS

Preparation of Cells

The isolation of the Syrian hamster melanoma line, HAB, has been described previously (1). HAB-B cells, having > 98.8% of the dT residues in DNA replaced by BrdU (1), were obtained by growing the HAB cells in monolayer culture in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 100 μ M hypoxanthine, 0.4 μ M aminopterin, and 10 μ M BrdU (growth medium). HAB-T cells, having dT but no BrdU in the DNA, were obtained by growing the HAB cells for over three months in growth medium containing 10 μ M dT instead of BrdU.

To obtain cells for DNA preparations, 10^5 cells were inoculated in 25 cm² tissue culture flasks and grown for 8 days in growth medium containing BrdU or dT in the case of HAB-B or HAB-T, respectively. Aliquots of 5×10^6 cells were collected by treatment with a solution of 0.05% trypsin and 0.02% EDTA and washed once with 5 ml of ice-cold phosphate-buffered saline (Ca⁺⁺-Mg⁺⁺-free). The pelleted aliquots of cells were immediately frozen in liquid nitrogen and stored until use for DNA preparation. Growth of the cells, as well as manipulations of the DNA, were carried out under lighting conditions which excluded wavelengths below 550 nm.

Isolation of DNA

Equal numbers of HAB-B and HAB-T cells were used to prepare comparable samples of B and T DNA. A modification of the procedure of Blin and Stafford (8) was employed. To a frozen pellet of 5×10^6 cells were added 100 μ l of K buffer (pH 8.0, 500 mM EDTA, 0.5% sarcosine) and 100 μ g of Proteinase K (EM Laboratories, Inc.). The mixture was placed immediately in a 37°C water bath and stirred with air bubbles from a micropipet. Since the mixture rapidly

became viscous, it was diluted by adding distilled water containing 0.5% sarcosine (400 μ l after 5 min and 500 μ l after 30 min). Incubation at 37°C was continued for a total of 3 hr.

The resultant solution was placed on ice and extracted with an equal volume of cold phenol saturated with K buffer. A total of three phenol extractions were made, after which the aqueous solution was extracted four times with ether to remove residual phenol. Distilled water was added to reduce the EDTA concentration to below 25 mM. DNA and RNA were then precipitated, without noticeable precipitation of EDTA, by adding 2.5 volumes of cold ethanol, gentle mixing and refrigeration at -20°C overnight.

The precipitate was pelleted by centrifugation for 30 min at 10,000 x g, rinsed once with ethanol, and dried under vacuum. To degrade RNA, the pellet was dissolved in 500 μ l of A buffer (pH 8, 50 mM Tris, 10 mM NaCl, 10 mM EDTA) and digested with 50 μ g of RNAase A (Worthington) at 37°C for 3 hr. Following 3 phenol extractions, 4 ether extractions, and the addition of 25 μ l of 3 M NaOAc, DNA was precipitated by the addition of 2.5 volumes of cold ethanol as before. After removal of the supernatant, the DNA pellet was redissolved in 300 μ l of distilled water. The B and T DNA samples so obtained were found by UV spectrophotometry and electrophoresis to be very similar in purity, molecular weight and concentration.

Restriction Digests

The restriction endonucleases used were obtained commercially (New England Biolabs: Alu I, Ava II, Bam HI, Bcl I, Hind III, Hinf I, Hpa I, Kpn I, Mbo I, Xba I; Bethesda Research Laboratories: Eco RII, Taq I; Calbiochem: Eco RI). The digestion of DNA was carried out using buffer conditions recommended in company catalogs. B and T DNA digestions were carried out under identical conditions. In each case, 1 μ g DNA was digested for 1 hr in a reaction volume of 20 μ l in a tightly capped polypropylene tube. The amounts of enzyme in the reaction mixture ranged from 1/16 unit to 10 units. The reactions were quenched by adding 2 μ l of K buffer (0.5 M EDTA) and freezing.

Electrophoresis

The DNA before and after restriction digestion was examined by electrophoresis on horizontal slab gels (16 cm x 20 cm) of 1.4% agarose in TBE buffer (pH 8.3, 90 mM Tris-Borate, 2.5 mM EDTA). Electrophoresis was carried out at room temperature for 14 - 16 hr at 5 mA or 2 volts/cm. DNA samples of 1 μ g were loaded in a volume of 25 μ l containing 5% sucrose and the dye markers, bromophenol blue and xylene cyanole. The gels were stained with ethidium bromide (1 mg/l) for 1 hr, rinsed for 15 min with distilled water, and photo-

graphed under UV illumination with Polaroid Type 55 P/N film.

Densitometry

Photographic negatives of electrophoretic gels were scanned with a Joyce-Loebl densitometer to determine the distribution of DNA as a function of migration distance. The densitometer curves were integrated with a Numonics graphic calculator.

The extent of cleavage was followed by measuring the fraction (f) of total DNA cleaved into fragments migrating ahead of the leading edge of the original band of uncleaved DNA. Integration of the densitometry curve from the position of the sample well to the bottom of the gel gave the total DNA intensity (I). Integration from the position of the leading edge of the original DNA band to the bottom of the gel gave the low molecular weight intensity (L). The fraction f was evaluated by our standard method (5), $f = (I/I_0 - I_0/I_0)/(1 - I_0/I_0)$, where the subscript o refers to undigested DNA.

Considering that intensity may be a non-linear function of DNA concentration, we also examined an alternative way of evaluating f . This was $f = (L - L_0)/(I_m - L_0)$, where m refers to maximally digested DNA. Had intensity been directly proportional to concentration, we would have found $I = I_0 = I_m$ and obtained the same f -values by the two methods. Instead we found $I_0 < I < I_m$ (5) and observed that f was likely overestimated by the standard method and underestimated by the alternative method. Nevertheless, upon plotting f_B/f_T vs. f_B , we found that both methods yielded essentially the same straight lines. Since the standard method consistently gave a better least-squares fit, this was the method of choice.

Kinetic Analysis

In our experiments, the DNA concentration is held constant while the enzyme concentration, $[E]$, is varied. The rate at which the DNA degraded fraction f appears with time t is related to $[E]$ by the equation,

$$\frac{df}{dt} = k[E] \frac{f_{\max} - f}{f_{\max}}$$

where k is a rate constant characteristic of each enzyme and $(f_{\max} - f)/f_{\max}$ is the fraction of DNA that remains undegraded at any time ($f_{\max}$ being the f -value for the limit digest).

Our aim is to evaluate k_B/k_T , the ratio of the rate constants for B and T DNA cleavage. An equation for this ratio can be found by integrating the expression for $df/(f_{\max} - f)$ for B and T DNA under conditions where the DNA concentration, the enzyme concentration, and the time of digestion are the same. The result is

$$\frac{k_B}{k_T} = \frac{\log(1 - f_B/f_{\max})}{\log(1 - f_T/f_{\max})} = \lim_{f_B \rightarrow 0} \frac{f_B}{f_T}.$$

Therefore, to evaluate k_B/k_T , we simply make a plot of f_B/f_T vs. f_B and extrapolate to $f_B = 0$. By extrapolating to $f_B/f_T = 1$, we also find $f_{\max}$.

RESULTS

Effect of BrdU Substitution on Restriction Endonuclease Activity

Using the experimental approach described for Hpa I and Mbo I (5), we have now examined the effect of BrdU substitution on eleven additional restriction endonucleases of different sequence specificity. The same preparations of unsubstituted (T) DNA and 100% BrdU substituted (B) DNA were used for all the nucleases. For each nuclease, the fractional degradation of B and T DNA was examined as a function of enzyme concentration by gel electrophoresis, fluorescent staining and densitometry. As described in Materials and Methods, a plot was made of f_B/f_T vs. f_B , where f_B and f_T represent the fractional degradation evaluated for B and T DNA, respectively, at each enzyme concentration. The results are shown in Fig. 1 for each of the eleven additional restriction endonucleases examined here. Except for Eco RI, the plots indicate a linear relationship, as found previously for Hpa I and Mbo I (5). Linear extrapolation of the value of f_B/f_T to $f_B = 0$ is used to determine k_B/k_T , the ratio of the rate constants for B and T DNA cleavage. Extrapolation of the line to $f_B/f_T = 1$ gives the maximum value ($f_{\max}$) of the fraction degraded, with the exception of Eco RI. The resultant k_B/k_T and $f_{\max}$ values for each nuclease are shown in Table I along with the values previously reported for Hpa I and Mbo I (5). The data are arranged according to the number of bases (N) in the recognition sequence for each nuclease.

Theoretical Formulation

In our previous paper (5), we suggested that the substitution of BrdU for dT exerts both an inhibitory and a stimulatory effect on the rate of DNA cleavage. We further suggested that the inhibitory effect results from steric hindrance by Br, whereas the stimulatory effect results from tighter binding to Br. These suggestions are consistent with known features of a Br atom and a CH₃ group covalently bound to an aromatic ring (Table II). Perpendicular to the covalent bond, Br and CH₃ have essentially the same size, as shown by their equivalent Van der Waals radii. However, parallel to the bond, Br is longer than CH₃ by about 0.37 Å, the difference between the covalent bond radii of Br (1.14 Å) and C (0.77 Å). Being longer, the Br atom in BrdU

Fig. 1. Results of plotting f_B/f_T vs. f_B for the groups of enzymes (a - d) listed in Table I, except for Hpa I and Mbo I. The results for the latter appear in a previous publication (5). Linear extrapolation of f_B/f_T to $f_B = 0$ for each enzyme yields the value of k_B/k_T shown in Table I. Extrapolation of f_B to $f_B/f_T = 1$ yields the corresponding value of f_{max} . The curvature seen in the data for Eco RI may be indicative of a small amount of Eco RI* activity (see text).

extends farther out into the wide groove of DNA than does the CH₃ group in dT. Thus, Br can hold the nuclease farther away from the DNA by steric hindrance (Fig. 2) and cause an inhibitory effect on DNA cleavage. On the other hand, the enzyme can bind more tightly to Br than to CH₃, since Br has the same Van der Waals radius as CH₃ and is somewhat more hydrophobic (Table II). A tighter binding would give rise to a stimulatory effect. To test these ideas, we have developed an empirical formula to describe the resultant inhibitory and stimulatory contributions to $\log k_B/k_T$ for 100% substitution.

TABLE I. Values of k_B/k_T and $f_{\max}$ found for restriction nucleases that cleave specific N-base sequences.

Nuclease	Base Sequence Cleaved	k_B/k_T	$f_{\max}$
(a) <u>N=6 ($N_T=1$)*</u>			
Bam HI	GGATCC [†]	0.50 ± 0.05	0.54
Kpn I	GGTACC [†]	0.40 ± 0.05	0.54
(b) <u>N=6 ($N_T=2$)*</u>			
Xba I	TCTAGA [†]	0.75 ± 0.05	0.77
Eco RI	GAATTC [†]	0.70 ± 0.10	—
Bcl I	TGATCA [†]	0.35 ± 0.05	0.77
Hind III	AAGCTT [†]	0.28 ± 0.04	0.77
Hpa I	GTTAAC [†]	$0.12 \pm 0.02\#$	0.74 [‡]
(c) <u>N=5</u>			
Eco RII	CC [†] AGG [†]	0.85 ± 0.10	0.94
Ava II	GG [†] TCC [†]	0.85 ± 0.10	0.94
Hinf I	GAXTC [†]	0.25 ± 0.05	0.97
(d) <u>N=4</u>			
Mbo I	GATC [†]	$5.0 \pm 1.0\#$	0.98 [‡]
Alu I	AGCT [†]	0.80 ± 0.10	0.98
Taq I	TGA [†]	0.42 ± 0.05	0.89 [§]

* N_T = No. of T's in sequence of N bases.

† X = A, C, G or T.

As determined previously (5).

§ Finding $f_{\max}$ lower for Taq I than for Mbo I and Alu I is in keeping with the well-known observation that GC has a low frequency of occurrence in mammalian DNA.TABLE II. Features of -Br and -CH₃ in aromatic molecules.

Feature	-Br	-CH ₃	Reference
Van der Waals radius	1.95 Å	2.0 Å	(11)
Covalent bond radius	1.14 Å	0.77 Å	(11)
Hydrophobicity	0.86	0.56	(12)

The formula, which contains two empirical parameters (α and β) for the inhibitory effect and a single parameter (σ) for the stimulatory effect, is as follows:

$$\log \frac{k_B}{k_T} = \frac{(N-N_T)}{N} \sum \left[\frac{\alpha}{p_a^n} + \frac{\beta}{p_b^n} \right] + \frac{N_T}{N} (\sigma)$$

where N = total number of bases in the cleavage site

N_T = number of T's in the cleavage site

p_a = position index for each T with respect to the cleavage point on the same strand

p_b = position index for each T with respect to the cleavage point on the opposite strand

n = exponent of the position index

To derive this formula we assume that a nuclease interacts with each T in the cleavage site by providing a hydrophobic pocket that fits the CH_3 group of T (Fig. 2). We also assume that Br fills the cavity and binds more tightly because of its greater hydrophobicity. On the grounds that each cavity is similar in kind, the enhanced binding is predicted to be independent of position in the cleavage site. The resultant stimulatory contribution to log

Fig. 2. Illustration of an enzyme cavity designed to fit a methyl group (CH_3) on DNA. When the cavity contains CH_3 , the enzyme is in close contact with DNA (upper diagram). When the cavity contains a bromine atom (Br) instead, the enzyme is held away from the DNA by steric hindrance (lower diagram).

TABLE III. Predicted inhibitory and stimulatory contributions to $\log k_B/k_T$ for each restriction nuclease.

Nuclease	Cleavage Site	Predicted Contributions	
		Inhibitory	Stimulatory
Bam HI	$\begin{array}{c} \uparrow \\ \text{GGATCC} \\ \text{CCTAGG} \end{array}$	$\frac{5}{6} \left(\frac{\alpha}{3^n} + \frac{\beta}{2^n} \right)$	$\frac{\sigma}{6}$
Kpn I	$\begin{array}{c} \uparrow \\ \text{GGTACC} \\ \text{CCATGG} \end{array}$	$\frac{5}{6} \left(\frac{\alpha}{3^n} + \frac{\beta}{2^n} \right)$	$\frac{\sigma}{6}$
Xba I	$\begin{array}{c} \uparrow \\ \text{TCTAGA} \\ \text{AGATCT} \end{array}$	$\frac{2}{3} \left(\alpha + \frac{\alpha}{2^n} + \frac{\beta}{3^n} + \frac{\beta}{5^n} \right)$	$\frac{\sigma}{3}$
Eco RI	$\begin{array}{c} \uparrow \\ \text{GAATTC} \\ \text{CTTAAG} \end{array}$	$\frac{2}{3} \left(\frac{\alpha}{3^n} + \frac{\alpha}{4^n} + \beta + \frac{\beta}{2^n} \right)$	$\frac{\sigma}{3}$
Bcl I	$\begin{array}{c} \uparrow \\ \text{TGATCA} \\ \text{ACTAGT} \end{array}$	$\frac{2}{3} \left(\alpha + \frac{\alpha}{3^n} + \frac{\beta}{2^n} + \frac{\beta}{5^n} \right)$	$\frac{\sigma}{3}$
Hind III	$\begin{array}{c} \uparrow \\ \text{AAGCTT} \\ \text{TTCGAA} \end{array}$	$\frac{2}{3} \left(\frac{\alpha}{4^n} + \frac{\alpha}{5^n} + 2\beta \right)$	$\frac{\sigma}{3}$
Hpa I	$\begin{array}{c} \uparrow \\ \text{GTTAAC} \\ \text{CAATTC} \end{array}$	$\frac{2}{3} \left(\alpha + \frac{\alpha}{2^n} + \beta + \frac{\beta}{2^n} \right)$	$\frac{\sigma}{3}$
Eco RII	$\begin{array}{c} \uparrow \\ \text{CCTGG} \\ \text{GGACC} \end{array}$	$\frac{9}{10} \left(\frac{1}{2} \left(\frac{\alpha}{3^n} \right) + \frac{1}{2} \left(\frac{\beta}{3^n} \right) \right)$	$\frac{\sigma}{10}$
Ava II	$\begin{array}{c} \uparrow \\ \text{GGTCC} \\ \text{CCAGG} \end{array}$	$\frac{9}{10} \left(\frac{1}{2} \left(\frac{\alpha}{2^n} \right) + \frac{1}{2} \left(\frac{\beta}{2^n} \right) \right)$	$\frac{\sigma}{10}$
Hinf I	$\begin{array}{c} \uparrow \\ \text{GAXTC} \\ \text{CTXAG} \end{array}$	$\frac{3}{4} \left(\frac{\alpha}{3^n} + \frac{1}{4} \left(\frac{\alpha}{2^n} \right) + \beta + \frac{1}{4} \left(\frac{\beta}{2^n} \right) \right)$	$\frac{\sigma}{4}$
Mbo I	$\begin{array}{c} \uparrow \\ \text{GATC} \\ \text{CTAG} \end{array}$	$\frac{3}{4} \left(\frac{\alpha}{3^n} + \frac{\beta}{2^n} \right)$	$\frac{\sigma}{4}$
Alu I	$\begin{array}{c} \uparrow \\ \text{AGCT} \\ \text{TCGA} \end{array}$	$\frac{3}{4} \left(\frac{\alpha}{2^n} + \frac{\beta}{2^n} \right)$	$\frac{\sigma}{4}$
Taq I	$\begin{array}{c} \uparrow \\ \text{TCGA} \\ \text{AGCT} \end{array}$	$\frac{3}{4} \left(\alpha + \frac{\beta}{3^n} \right)$	$\frac{\sigma}{4}$

k_B/k_T is simply proportional to the mole-fraction of T's in the cleavage site and can be expressed as $\sigma N_T/N$, where σ has a similar positive value for all nucleases under identical conditions.

The inhibitory contribution, on the other hand, depends on the position

of each T in the cleavage site. Since Br extends farther into the wide groove than CH₃, it can sterically hinder the binding of nuclease to unsubstituted bases and to the points of cleavage. The degree of inhibition is related to the mole-fraction of unsubstituted bases, $(N-N_T)/N$, and inversely related to the distance from the point of substitution to the point of cleavage. The relative distance can be estimated in terms of a position index, P. If T lies in the base pair immediately adjacent to the point of cleavage, P = 1; if T lies one base pair farther away, P = 2, etc. The decline with distance can be described approximately as $1/P^n$, where n is the exponent of distance. Each T has two values of the position index, P_a and P_b, referring to the cleavage point on the same strand as T and opposite strand, respectively. The inhibitory contribution from each substituted T can be described in terms of the average for the two cleavage points, $\frac{1}{2}(\alpha/P_a^n + \beta/P_b^n)$, where α and β have similar negative values in all nucleases. Thus the total inhibitory contribution is the product of $(N-N_T)/N$ and $\sum \frac{1}{2}(\alpha/P_a^n + \beta/P_b^n)$, where the summation is carried out for all substituted T's on both strands in the cleavage site.

In Table III are shown the predicted inhibitory and stimulatory contributions to $\log k_B/k_T$ for each of the nucleases studied. The evaluation of α and β is made simple by the fact that certain enzymes have terms in common. The pair Eco RI and Hpa I have the same terms in β and the same terms in σ . Thus, for any given value of n, α can be determined from the observed difference in the $\log k_B/k_T$ values for Eco RI and Hpa I. Similarly, since Xba I and Hpa I have the same terms in α and the same terms in σ , the value of β can be determined from the difference in the $\log k_B/k_T$ values of Xba I and Hpa I.

Having so evaluated α and β , we can determine the constancy of σ by a least-square fit to the data for all the enzymes. For n = 1, we find $\alpha = -1.25$, $\beta = -1.23$, and $\sigma = 4.04 \pm 0.87$. The corresponding standard deviation between predicted and observed values of $\log k_B/k_T$ is ± 0.202 . For n values above and below n = 1, the standard deviation increases (Table IV). Thus, the formula fits the data best with n = 1, indicating that the inhibitory effect is inversely proportional to the first power of the distance from the cleavage point. In addition, the data indicate that α and β are approximately equal.

A possible reason for discrepancy between theory and experiment is the difference in ionic strength (μ) in the reaction buffers used. The values of μ ranged from 0.03 for Kpn I to 0.175 for Mbo I. In order to analyze the possible influence of ionic strength, we examined the difference (Δ) between the

TABLE IV. Variation in the standard deviation, between observed and predicted $\log k_B/k_T$, as the value of the exponent n in Table II is varied.

n	Standard Deviation
1/3	0.350
1/2	0.228
1	0.202
2	0.269
3	0.320
∞	0.395

observed and predicted $\log k_B/k_T$ values as a function of μ . As shown in Table V, there is a strong correlation between Δ and μ , with a correlation coefficient of 0.80. On the other hand, the observed and predicted $\log k_B/k_T$ values themselves show less correlation with μ , giving correlation coefficients of 0.6 and 0.2, respectively.

We also considered the possibility that the inhibitory and stimulatory parameters vary with μ in a simple manner. To explore this possibility, we set $n = 1$ and $\beta = \alpha$, and empirically evaluated α and σ for each group of enzymes examined under similar ionic strength. The results suggest the simple relations, $\alpha = \beta = -4\sqrt{\mu}$ and $\sigma = 40\mu$. In terms of μ , the formula then becomes

$$\log \frac{k_B}{k_T} = -4\sqrt{\mu} \left(\frac{N-N_T}{N} \right) \left[\frac{1}{2} \left(\frac{1}{p_a} + \frac{1}{p_b} \right) + 40\mu \left(\frac{N_T}{N} \right) \right].$$

This formula fits the data with a standard deviation of 0.112, which is approximately one-half of the standard deviation found without taking into account the ionic strength of the buffer.

DISCUSSION

We have developed a model to explain the effect of substituting BrdU for dT in DNA on the rate of sequence-specific cleavage. The model predicts that substitution causes an inhibitory effect which varies with position in the cleavage site and a stimulatory effect which is independent of position. The empirical formula for $\log k_B/k_T$ derived from the model agrees quite well with the results obtained for thirteen different restriction endonucleases assayed

TABLE V. Listing of nucleases showing that the difference (Δ) between observed and predicted $\log k_B/k_T$ is strongly correlated with the ionic strength (μ) of the reaction buffer.

Nuclease	μ^*	$\Delta^\dagger$
Mbo I	0.175	0.47
Xba I	0.170	0.22
Eco RI	0.150	0.22
Hind III	0.090	0.12
Ava II	0.075	0.09
Hinf I	0.075	-0.14
Bam HI	0.075	-0.11
Alu I	0.075	-0.17
Eco RII	0.075	-0.10
Hpa I	0.060	0.22
Bcl I	0.060	-0.11
Taq I	0.030	-0.14
Kpn I	0.030	-0.21

Coefficient of correlation with μ 0.80

* As determined from the molar concentration (C) of added salt (NaCl or KCl), the HCl salt of trizma base, and MgCl₂, by the relationship, $\mu = C_{\text{salt}} + C_{\text{Tris} \cdot \text{HCl}} + 3C_{\text{MgCl}_2}$.

† Difference between $\log k_B/k_T$ found experimentally for each nuclease (Table I) and the value predicted from the inhibitory and stimulatory contributions (Table II) with $n = 1$, $\alpha = \beta = -1.24$, and $\sigma = 4.04$.

under somewhat different buffer conditions. When the ionic strengths of the buffers are taken into account, an even better agreement between prediction and result is obtained (Fig. 3). Of course, we have not ruled out the possibility that some of the effects observed may depend on individual enzyme properties not related to sequence specificity. Comparisons of isoschizomers such as Mbo I and Sau 3A may help reveal whether or not this is the case.

The formula can be used to predict the effect of BrdU substitution on any restriction endonuclease whose specific cleavage sequence is known. From our results it appears that the formula holds equally well for nucleases which recognize sequences containing from 4 to 6 base pairs. In the case of Eco RI, the plot of f_B/f_T vs. f_B shows a distinct curvature not seen with the other enzymes (Fig. 1 (b)). On the basis of the formula, we attribute the curvature

Fig. 3. Comparison of the observed values of $\log k_B/k_T$ with those predicted by our formula with $n = 1$, $\alpha = \beta = -4\sqrt{\mu}$, and $\sigma = 40 \mu$.

to the presence of a small amount of Eco RI* activity. Considering that the specificity of Eco RI* cleavage is ^+AATT (11), we predict k_B/k_T to be in the range 5 - 12 at the ionic strength used (0.15). Thus, f_B/f_T values greater than 1 are expected to be found, as in the case of Mbo I (5). We do, in fact, see $f_B/f_T > 1$ at large values of f_B in the data for Eco RI (Fig. 1 (b)). This would be consistent with there being a slow Eco RI* activity, which is greater for B DNA than for T DNA and which continues after all of the Eco RI sites have been cleaved.

The formula is based entirely on considerations of the effects of BrdU within the confines of the specific sequence recognized by each nuclease. Studies with λ DNA show that cleavage rates can vary from site to site in the DNA molecule (12). Presumably this has to do with interactions between the nuclease and portions of the DNA outside of the specific recognition sequence. The fact that we are able to obtain a good fit to our data for a large number of nucleases may be a consequence of our having used mammalian genome in which there is a great variety of environments around the cleavage sites. In such DNA, interactions outside the cleavage site might be expected to average out to be roughly the same from one nuclease to another. Had we used a simpler

DNA such as λ , it is less likely that we would have been able to explain our results in terms of a general formula. In comparing our results with those of Berkner and Folk (4) on λ DNA, we are in agreement in the k_B/k_T values for Eco RI and Hind III, but not for Bam HI. Since there are few Bam HI sites on λ , it is possible that there are some unusual outside interactions which give rise to the discrepancy.

Our formulation extends and modifies the suggestions made in our previous paper (5). We had previously proposed that the inhibitory effect falls off extremely rapidly with distance and is important only when the base is immediately adjacent to the cleavage point. If this were the case, the exponent n would be large. As seen in Table IV, this idea is not borne out by the data. Instead, the inhibitory effect appears to fall off as the first power of the distance, i.e., $n = 1$. Furthermore, we had suggested that the inhibitory parameter α might be greater than β . We now see that α and β are approximately equal in magnitude, so that only one inhibitory parameter need be used.

Furthermore, the data presented in this paper suggest that the inhibitory and stimulatory parameters are sensitive to the ionic strength (μ) of the reaction buffer. Over the range of buffers used, the simple relationships $\alpha = \beta = -4\sqrt{\mu}$ and $\sigma = 40\mu$ appear to hold. However, studies of salt effects on individual enzymes will be needed to verify these relationships.

It remains to be shown that our theoretical formulation is applicable to

TABLE VI. Evidence for stimulatory and inhibitory contributions to $\log t_B/t_T$ in lac repressor-operator binding.

Neighborhood of T*	Average value of $\log t_B/t_T$ for single substitution of B for T		
	Predicted contributions [†]	Observed with SQ repressor [‡]	Observed with QX86 repressor [‡]
RTR	σ'	0.23 ± 0.04	0.15 ± 0.06
YTR	$\sigma' + \alpha'$	0.06 ± 0.03	0.07 ± 0.06
RTY	$\sigma' + \alpha'$	-0.07 ± 0.05	-0.03 ± 0.04
YTY	$\sigma' + 2\alpha'$	-0.13 ± 0.04	-0.12 ± 0.09

* R = purine; Y = pyrimidine.

[†] σ' and α' are the stimulatory and inhibitory parameters, respectively, in the predicted formula, $\log t_B/t_T = \sigma'N_B + \alpha'N_Y$ where N_B = no. of T's substituted by B's and N_Y = no. of unsubstituted pyrimidines in the immediate neighborhood of the substituted T's.

[‡] Evaluated for 0.05M ionic strength from the data in Table 6, ref. (13).

single substitutions of BrdU for dT within the cleavage site. Although studies of individual substitutions have yet to be made with any endonuclease, they have been made with lac repressor (13). In this case, the effects on binding can be evaluated in terms of the half-life of repressor-operator complex.

Our model predicts that repressor binds more tightly to Br than to CH₃ (on T) but less tightly to groups on neighboring bases because of the steric hindrance by Br. Therefore, the character of the bases flanking the substituted T should be most important in determining the magnitude of the inhibitory effect. Since pyrimidines (Y) are more hydrophobic in the wide groove than purines (R), they have a stronger hydrophobic contribution to binding. Considering that non-polar interactions decline with distance more rapidly than do polar interactions, we expect steric hindrance to weaken the binding to Y more than to R. On this basis, we predict the inhibitory contribution to be roughly proportional to the number of adjacent pyrimidines, N_Y. Accordingly, if t_B and t_T are the half-lives for repressor binding to substituted and unsubstituted operator, respectively, our prediction for single substitution is that

$$\log \frac{t_B}{t_T} = \sigma' + \alpha' N_Y$$

where σ' and α' are positive and negative constants, respectively, analogous to σ/N and α/N for restriction endonucleases.

The prediction can be compared with experimental data in the literature (13) by taking the average value of $\log t_B/t_T$ for each kind of neighborhood of T in the operator site, namely RTR, YTR, RTY, and YTY. In Table VI are shown average values for two kinds of repressor (SQ and QX86) compared with the predictions made by our simple equation. These data reveal a trend consistent with our predictions. Substitution in RTR sites has a net stimulatory effect ($\sigma' = 0.23$ and 0.15 for SQ and QX86, respectively). Substitution in YTY sites has a net inhibitory effect ($\alpha' = -0.18$ and -0.14 , respectively). For the intermediate types of sites, RTY and YTR, the stimulatory and inhibitory contributions almost cancel out. Thus it appears likely that our general concept of inhibitory and stimulatory effects, based on steric and hydrophobic considerations, holds approximately for repressors as well as for restriction endonucleases.

ACKNOWLEDGEMENTS

This work was supported by the Kammerman Fund and by NIH Biomedical Research Support Grant 5 S07 RR07012. We thank Sarah Wright and Amy Wong for

their help in preparation of the manuscript.

REFERENCES

1. Bick, M. D., and Davidson, R. L. (1974). *Proc. Natl. Acad. Sci. U.S.A.* 71, 2082-2086.
2. Lin, S. Y., and Riggs, A. D. (1972). *Proc. Natl. Acad. Sci. U.S.A.* 69, 2574-2576.
3. Marchionni, M. A., and Roufa, D. J. (1978). *J. Biol. Chem.* 253, 9075-9081.
4. Berkner, K. L., and Folk, W. R. (1979). *J. Biol. Chem.* 254, 2551-2560.
5. Petruska, J., and Horn, D. (1980). *Biochem. Biophys. Res. Commun.* 96, 1317-1324.
6. Rutter, W. J., Pictet, R. L., and Morris, P. W. (1973). *Ann. Rev. Biochem.* 42, 601-646.
7. Smith, H. O. (1979). *Science* 205, 455-462.
8. Blin, N., and Stafford, D. W. (1976). *Nucl. Acids Res.* 3, 2303-2309.
9. Pauling, L. (1948) *Nature of the Chemical Bond*, 2nd edn., pp. 160-193, Cornell University Press, Ithaca.
10. Rekker, R. F. (1977) *The Hydrophobic Fragmental Constant*, pp. 25-37, Elsevier Scientific Publishing Company, New York.
11. Polisky, B., Green, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M., and Boyer, H. W. (1975). *Proc. Natl. Acad. Sci. U.S.A.* 72, 3310-3314.
12. Thomas, M. and Davis, R. W. (1975). *J. Mol. Biol.* 91, 315-328.
13. Goeddel, D. V., Yansura, D. G., Winston, C., and Caruthers, M. H. (1978) *J. Mol. Biol.* 123, 661-687.