Kinetics of Promoter Search by *Escherichia coli* RNA Polymerase

EFFECTS OF MONOVALENT AND DIVALENT CATIONS AND TEMPERATURE

(Received for publication, September 8, 1987)

Paul T. Singer and Cheng-Wen Wu‡

From the Department of Pharmacological Sciences, State University of New York at Stony Brook, Stony Brook, New York 11794

The rapid mixing/photocross-linking technique developed in our laboratory has been employed in the study of the mechanism of promoter binding by *Escherichia coli* RNA polymerase (RPase). We have previously reported on the quantitation of the one-dimensional diffusion coefficient ($D_1$) for RPase along the DNA template (Singer, P. T., and Wu, C.-W. (1987) J. Biol. Chem. 262, 14178–14189). In this paper, we describe the effects of salt concentration and temperature on the kinetics of promoter search by RPase using plasmid pH1319 DNA, which contains the A2 early promoter from bacteriophage T7, as template. Over a range of KCl concentrations from 25 to 200 mM, the apparent bimolecular rate constant ($k_a$) for the association of RPase with the A2 promoter on this DNA template varied approximately 2-fold, achieving a maximal value between 100 and 125 mM KCl. More significantly, the transient distribution of RPase among nonspecific DNA binding sites changed markedly as a function of salt concentration, indicative of gross changes in the average number of base pairs covered by sliding during a nonspecific lifetime. Using the mathematical treatment outlined in our earlier report, the nonspecific dissociation rate constant ($k_{off}$) was calculated from the binding curves for the nonspecific as well as promoter-containing DNA. The observed variations in $k_a$ as a function of monovalent cation concentration ([M$^+$]) were due primarily to changes in $k_{off}$, as $D_1$ was found to be essentially independent of [M$^+$]. Interestingly, $D_1$ decreased by one-third as the concentration of magnesium was lowered from 10 to 1 mM. In addition, the dependence of $k_{off}$ (and consequently the nonspecific equilibrium association constant, $K_{eq}$) on [M$^+$] agreed qualitatively with the results of deHaseth et al. (deHaseth, P. L., Lohman, T. M., Burgess, R. R., and Record, M. T., Jr. (1977) Biochemistry 17, 1612–1622), though we consistently measured a weaker $K_{eq}$. The association rate constant was also measured between 4 and 37 °C, and was found to vary 2-fold over that range. An activation energy for the bimolecular association of RPase to the A2 promoter was calculated to be 2.2 ± 0.4 kcal/mol, while the activation energy for one-dimensional diffusion was 4.7 ± 0.8 kcal/mol.

The elucidation of the molecular mechanism of promoter search by *Escherichia coli* RNA polymerase (RPase) has been the subject of research for many years, in that location of promoter sequences by this enzyme is the initial step in gene transcription. As such, it has the potential to serve as a means of regulation of the expression of a wide variety of genes in the bacterium. That promoter search is not a simple bimolecular binding of RPase to promoter has been postulated based on the anomalously high rates measured for binding of RPase to a variety of promoters (1, 2). In order to explain these observations, it has been proposed that RPase (R) binds initially to nonspecific sites on the DNA (D) via a bimolecular mechanism, and subsequently transfers intramolecularly to the promoter (P):

$$R + D + P \overset{k_1}{\rightarrow} R.D + P \overset{k_2}{\rightarrow} R.P + D \overset{k_{off}}{\rightarrow}$$

One mechanism proposed to account for the speed of the second step is the intramolecular transfer of RPase from nonspecific to specific sites via one-dimensional diffusion (sliding) along the DNA molecule. While other facilitating mechanisms, such as “hopping” or “intersegment transfer” are possible (see Ref. 3), sliding is the most efficient on theoretical grounds (by orders of magnitude) and, if it occurs, should therefore mask any effects of the other mechanisms.

In previous studies, we have demonstrated both qualitatively and quantitatively that RPase locates promoter sequences via linear diffusion, on both linear plasmid and circular plasmid DNA (4) and the relaxed circular plasmid pAR1319 DNA (5), which contains the T7 A2 early promoter. The primary evidence was the correlation between the kinetics of the transient binding of RPase to nonspecific sites and the distance between those sites and the promoter. Evidence for linear diffusion as a factor in specific-site binding has also been demonstrated in a variety of other experimental systems including lac repressor/operator (6) and various restriction endonucleases (7, 8). Furthermore, it has been suggested (9) that eukaryotic enhancer sequences may exert their long-range action on promoters via a mechanism that requires the sliding of one or more regulatory proteins along the DNA between the two binding sites.

An earlier report from this laboratory measured the intradomain nonspecific association rate constant for RPase to pAR1319 DNA ($k_{off}$) to be 4.6 × 10$^4$ M$^{-1}$ s$^{-1}$ (5). Furthermore, the rapid mixing/photocross-linking technique developed in our laboratory (10, 11) allows us to investigate the kinetics of interaction between RPase and pAR1319 DNA by monitoring

---

*This work was supported by research grants from the National Institutes of Health (GM 2969), the National Science Foundation (PCM 8009813), and the American Cancer Society (NP 309). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom all correspondence should be addressed.

1 The abbreviations used are: RPase, *E. coli* DNA-dependent RNA polymerase; bp, base pair; AMV-RT, avian myeloblastoma virus DNA-dependent RNA polymerase (reverse transcriptase).
the time-dependent change in RPase concentration along the circular plasmid DNA molecule, in addition to RPase accumulation at the A2 promoter. Analysis of the transient RPase occupancy of nonspecific DNA binding sites using a mathematical model of promoter search makes it possible to fit the individual rate constants $D_i$ (one-dimensional diffusion coefficient) and $k_{off}$ (nonspecific dissociation rate constant) to the kinetic data (5). The rate constants, $k_{on}$ and $D_i$, were determined to be $0.5 \text{ s}^{-1}$ and $1.5 \times 10^{-9} \text{ cm}^2\text{s}^{-1}$, respectively, at $37 \degree C$ in 50 mM KCl and 10 mM MgCl$_2$ (5). This latter figure is approximately 4-fold smaller than the theoretically calculated limit for a rotating sphere moving along the major groove of the DNA molecule (12).

While it has been well established that the rate of specific-site location is dependent upon a variety of external parameters, such as temperature and salt (both monovalent and divalent cation) concentration, it is of considerable importance to determine exactly which steps in the overall binding scheme are perturbed by changes in these parameters. To this end, we report here on the salt and temperature dependence of the bimolecular rate constant for RPase dissociation from nonspecific DNA binding sites ($k_{off}$) and the one-dimensional diffusion coefficient for RPase along the DNA template ($D_i$).

**Mathematical Model of Promoter Search**—In this model, which is described in greater detail in an earlier report (5), the plasmid DNA molecule is described as a circular array of "compartments" ($2n + 1$ of which, each containing multiple intrinsic binding sites) which together define a DNA "domain." RPase diffuses into and out of the domain at the diffusion controlled limit, and bimolecularly binds to ($k_{on}$) and dissociates from ($k_{off}$) the DNA molecule. It is important to note that these two rate constants describe an intradomain phenomenon, rather than the more usual bulk binding reaction portrayed in Equation 1. Once bound to the DNA the protein is free to diffuse linearly between adjacent compartments (i.e. sites); this intramolecular transfer is described by the one-dimensional diffusion coefficient, $D_i$, with an appropriate change of units from cm$^2$ s$^{-1}$ to either bp$^2$ s$^{-1}$ or compartments$^{-1}$. The transient concentrations of RPase in the compartments were represented by a family of differential rate equations incorporating terms accounting for the initial random binding of RPase to pAR1319, dissociation from nonspecific DNA, and linear diffusion of RPase along the DNA molecule.

$$\frac{d[R_n]}{dt} = D_i([R_{n-1}] + [R_{n+1}] - 2[R_n]) + S_k_{on}[R_0][D_0] - k_{off}[R_n] \tag{2}$$

where $[R_0], [R_{n+1}]$, and $[R_{n-1}]$ represent the concentrations of RPase in the $n$th compartment and its two nearest neighbors $[R_0]$ and $[D_0]$ are the intradomain concentrations of RPase and DNA base pairs, respectively, and $S_i$ is the compartment size in base pairs. Due to the symmetry of the circular DNA molecule, $[R_n]$ and $[R_{n-1}]$ are always identical, and are therefore represented by the same rate equation. Since the A2 promoter is effectively an infinite sink for RPase, $[R_0]$ is forced to zero wherever it appears in any of the rate equations.

**Experimental Procedures**

Materials—E. coli MRE 600 cells were purchased from Grain Processing, Inc. Unlabeled deoxyoligonucleotides triphosphates were obtained from Pharmacia LKB Biotechnology Inc. and [a-32P]dGTP was purchased from Amersham Corp. Avian myeloblastosis virus reverse transcriptase (AMV-RT) was from Life Sciences. Restriction endonucleases AvAI and HindIII were obtained from New England Biolabs. Proteinase K was from Boehringer Mannheim, and HeLa cell toposomerase I was the generous gift of Dr. Leroy Liu, The Johns Hopkins University Medical Center. Carrier E. coli tRNA was purchased from Sigma. All other biochemicals were of the highest purity commercially obtainable.

Preparation of RPase and Plasmid DNA—RPase was purified from E. coli MRE 600 cells as described previously (5), and was between 60 and 85% active as assayed by the method of Chamberlin et al. (13). Plasmid pAR1319, which was provided by Dr. William Studier and Dr. John Rosenberg at the Brookhaven National Laboratory, is a pBR322 derivative containing the Escherichia coli A2 early promoter and terminator sequences (Fig. 1). Relaxed plasmid DNA was prepared using standard techniques as described previously (5).

Rapid Mixing/Photocross-linking—The modifications to the Durran rapid mixing/photocross-linking apparatus were described in an earlier report from this laboratory (10). Using this apparatus, solutions of RPase and pAR1319 DNA in binding buffer (10 mM Tris-

HCl, pH 7.9, containing varying amounts of MgCl$_2$ and KCl or NaCl) were rapidly mixed (within 20 ms) into a UV-transparent quartz tube containing the noted final concentrations of RPase and pAR1319 DNA. Prior to mixing, both solutions were pre-equilibrated at the desired experimental temperature by submersion of the drive syringes in a circulating water bath. In addition, the quartz mixing tube was maintained at the same temperature as the syrings by thermal contact with a temperature-regulated aluminum block. At the time of mixing, the solution was exposed to a short (10 µs) pulse of high intensity UV light provided by a xenon arc lamp pulsed at 10 kW, which results in the formation of a covalent linkage between DNA and RPase.

Isolation and Analysis of Covalently Cross-linked DNA-RPase Complexes—Following UV-induced cross-linking as described above, the plasmid DNA was digested with AvAI and HindIII into nine fragments ranging in size from 42 to 1746 base pairs. The covalently cross-linked RPase-DNA complexes were selectively precipitated using sodium dodecyl sulfate as described previously (5), concentrated by ethanol precipitation, and the cross-linked RPase removed by exhaustive digestion with Proteinase K. The samples were then end-labeled with [a-32P]dGTP (20 µM) using AMV-RT. In order to standardize the labeling reaction, an identical aliquot of DNA was labeled in parallel with each experimental trial in order to account for the variations of incorporation of different lots of AMV-RT and radioactive nucleotides, thereby allowing comparison of results from one trial to the next. The seven largest fragments were separated by electrophoresis on a 25-cm 2.2% horizontal agarose gel for 16 h at 2.5 V/cm; the two smallest fragments, which were found to yield no useful information (5), were not quantitated. The radioactive bands were cut out of the gel, lyophilized, and counted using a Beckman LS-230 liquid scintillation counter.

**Materials and Methods**

1. **Preparation of RPase and Plasmid DNA**
   - RPase was purified from E. coli MRE 600 cells as described previously (5), and was between 60 and 85% active as assayed by the method of Chamberlin et al. (13). Plasmid pAR1319, which was provided by Dr. William Studier and Dr. John Rosenberg at the Brookhaven National Laboratory, is a pBR322 derivative containing the Escherichia coli A2 early promoter and terminator sequences (Fig. 1). Relaxed plasmid DNA was prepared using standard techniques as described previously (5).

2. **Rapid Mixing/Photocross-linking**
   - The modifications to the Durran rapid mixing/photocross-linking apparatus were described in an earlier report from this laboratory (10). Using this apparatus, solutions of RPase and pAR1319 DNA in binding buffer (10 mM Tris-

3. **Isolation and Analysis of Covalently Cross-linked DNA-RPase Complexes**
   - Following UV-induced cross-linking as described above, the plasmid DNA was digested with AvAI and HindIII into nine fragments ranging in size from 42 to 1746 base pairs. The covalently cross-linked RPase-DNA complexes were selectively precipitated using sodium dodecyl sulfate as described previously (5), concentrated by ethanol precipitation, and the cross-linked RPase removed by exhaustive digestion with Proteinase K. The samples were then end-labeled with [a-32P]dGTP (20 µM) using AMV-RT. In order to standardize the labeling reaction, an identical aliquot of DNA was labeled in parallel with each experimental trial in order to account for the variations of incorporation of different lots of AMV-RT and radioactive nucleotides, thereby allowing comparison of results from one trial to the next. The seven largest fragments were separated by electrophoresis on a 25-cm 2.2% horizontal agarose gel for 16 h at 2.5 V/cm; the two smallest fragments, which were found to yield no useful information (5), were not quantitated. The radioactive bands were cut out of the gel, lyophilized, and counted using a Beckman LS-230 liquid scintillation counter.

4. **Mathematical Model of Promoter Search**
   - In this model, which is described in greater detail in an earlier report (5), the plasmid DNA molecule is described as a circular array of "compartments" ($2n + 1$ of which, each containing multiple intrinsic binding sites) which together define a DNA "domain." RPase diffuses into and out of the domain at the diffusion controlled limit, and bimolecularly binds to ($k_{on}$) and dissociates from ($k_{off}$) the DNA molecule. It is important to note that these two rate constants describe an intradomain phenomenon, rather than the more usual bulk binding reaction portrayed in Equation 1. Once bound to the DNA the protein is free to diffuse linearly between adjacent compartments (i.e. sites); this intramolecular transfer is described by the one-dimensional diffusion coefficient, $D_i$, with an appropriate change of units from cm$^2$ s$^{-1}$ to either bp$^2$ s$^{-1}$ or compartments$^{-1}$. The transient concentrations of RPase in the compartments were represented by a family of differential rate equations incorporating terms accounting for the initial random binding of RPase to pAR1319, dissociation from nonspecific DNA, and linear diffusion of RPase along the DNA molecule.

   $$\frac{d[R_n]}{dt} = D_i([R_{n-1}] + [R_{n+1}] - 2[R_n]) + S_k_{on}[R_0][D_0] - k_{off}[R_n] \tag{2}$$

where $[R_0], [R_{n+1}]$, and $[R_{n-1}]$ represent the concentrations of RPase in the $n$th compartment and its two nearest neighbors $[R_0]$ and $[D_0]$ are the intradomain concentrations of RPase and DNA base pairs, respectively, and $S_i$ is the compartment size in base pairs. Due to the symmetry of the circular DNA molecule, $[R_n]$ and $[R_{n-1}]$ are always identical, and are therefore represented by the same rate equation. Since the A2 promoter is effectively an infinite sink for RPase, $[R_0]$ is forced to zero wherever it appears in any of the rate equations.

5. **Measurement of the Sliding Length**
   - The data was analyzed using a compartment size of 10 bp, because it was observed that a smaller compartment resulted in exactly the same results while requiring considerably more computation time. In some cases, however, the mean sliding length was insufficient short as to require a smaller compartment size for accurate fitting of the data. It was empirically determined that if the mean sliding length was less than approximately 8-fold longer than the compartment size, the results varied measurably with the choice of compartment size. Consequently, the data was refit in those cases using a compartment size of 10 bp.

6. **Computer Analyses**
   - The family of differential equations was solved numerically using a custom program written in Pascal, which employed the Runge-Kutta-Fehlberg algorithm (14). A second program, based on the Marquardt algorithm for non-linear, least squares analysis (15), used the numerically calculated functions to fit the important rate constants to the experimental data. These programs
were compiled and run on either an IBM-PC/AT microcomputer equipped with a 14 MHz 80287 math co-processor/daughterboard from Ariel Corp. or an IBM 4361 mainframe.

RESULTS

The rapid mixing/photocross-linking technique employed in this study yields data representing the occupancy by RNA polymerase of both specific and nonspecific sites along the DNA molecule at various time points immediately following the mixing of the two species (Fig. 2). As we have documented previously (5), there is sufficient information available from the six nonspecific curves to allow the two important rate constants, $k_{on}$ and $D_0$, to be fit to the data, rather than requiring that one be approximated in order to calculate the other. There are two types of data in these curves: the difference in kinetics of the RNA polymerase occupancy of the nonspecific fragments, and the total amount of nonspecific binding observed. In cases where the mean sliding distance ($L_s$) is long compared to fragment size, the individual bands are well-resolved, with respect to the standard error of the measurements (5). In cases where $L_s$ is short, the amount of nonspecific binding allows us to distinguish between the case where $k_{on}$ is fast and $D_0$ slow, and the case where the reverse is true, since only $k_{on}$ affects the amount of RNA polymerase bound nonspecifically to the DNA, while the distribution of bound protein is a complicated function of both $D_0$ and $k_{on}$. Since the quantitation of the transient occupancy by RNA polymerase of DNA binding sites involves cleavage of the pAR1319 DNA by restriction endonucleases, resolution is limited by the size of the resulting DNA fragments. Due to the great variation in size of the fragments (222 to 1746 bp), representation of the data as density of RNA polymerase (with units of counts/min/bp) facilitates comparison between fragments. As shown in Fig. 2, data plotted in this manner clearly demonstrates the random distribution of RNA polymerase along the pAR1319 molecule at early time points (<0.25 s), and the subsequent accumulation of most of the protein on Fragment B, which contains the A2 promoter. The specific association rate constant, $k_a$, was determined from the slope of the best-fit straight line through a semilogarithmic plot of the percent RNA polymerase associated with Fragment B, taking into account the initial nonspecific binding as described previously (5). The nonspecific association rate constant, $k_{on}$, was found to be $4.6 \times 10^4$ M$^{-1}$ s$^{-1}$, based on the assumption that the intradomain concentration of base pairs is 180 $\mu$M, as would be the case for a molecule of 4800 bp constrained within a sphere defined by a radius of gyration of 0.22 $\mu$m. The rate constants describing the two most important steps involved in promoter search (i.e. nonspecific dissociation, $k_{off}$, and one-dimensional diffusion, $D_0$) were calculated by fitting the experimental data for all seven fragments, using the mathematical model and the numerical methods described under "Experimental Procedures."

Effect of Monovalent Cation Concentration—Prior to rapid mixing, both the RNA polymerase and pAR1319 DNA were preincubated in binding buffer containing varying concentrations of KCl over the range 25–200 mM, while the concentration of MgCl$_2$ was fixed at 10 mM. All experiments were performed at 37 °C. Fig. 2 shows the RNA polymerase density for each of the DNA fragments plotted against time for three different salt concentrations (25, 50, and 150 mM), demonstrating the measurable differences in the kinetics of binding to both specific and nonspecific sites along the pAR1319 DNA molecule under varying solution conditions. In addition to the modest changes observed in the rate of association of RNA polymerase and the A2 promoter ($k_a$), the relative kinetics of RNA polymerase occupancy of the nonspecific fragments was strongly dependent on salt concentration. At low salt concentrations (Fig. 2A), the curves corresponding to the nonspecific fragments displayed markedly different kinetics, while at high salt concentrations (Fig. 2C) these curves were virtually superimposable. This phenomenon is due primarily to changes in $k_{on}$, which in turn affects the

![Figure 2](image-url)

**Fig. 2.** The kinetics of promoter search by RNA polymerase on pAR1319 DNA. RNA polymerase holoenzyme (3.2 nM, final concentration) and pAR1319 DNA (16 nM, final concentration) were rapid-mixed and UV cross-linked at the indicated times following mixing, as described under "Experimental Procedures." Each data point represents an average of at least three trials, performed in binding buffer containing 10 mM MgCl$_2$, 37 °C, at a KCl concentration of 25 mM (A) 100 mM (B) and 150 mM (C). The seven curves in each panel represent the RNA polymerase occupancy of: O, Fragment A; A, Fragment B; B, Fragment C; C, Fragment D; D, Fragment E; E, Fragment F; F, Fragment G. The error bars represent the standard errors for a minimum of three trials at each data point.

**TABLE I**

| [MgCl$_2$] [mM] | [KCl] [mM] | $k_{on} \times 10^{14}$ [s$^{-1}$] | $k_{off} \times 10^4$ [s$^{-1}$] | $k_m \times 10^4$ [s$^{-1}$] | $k_a \times 10^3$ [s$^{-1}$] | $D_0 \times 10^9$ [cm$^2$ s$^{-1}$] | $I_n$ [bp] |
|-----------------|----------|-------------------------------|-----------------------------|----------------|----------------|----------------|---------|
| 10              | 25       | 3.9                           | 0.1                         | 9.6            | 2.3            | 1.6            | 3,700   |
| 10              | 50       | 4.6                           | 0.3                         | 5.4            | 2.4            | 1.5            | 2,300   |
| 10              | 75       | 5.2                           | 1.3                         | 5.8            | 3.5            | 1.5            | 1,000   |
| 10              | 100      | 4.3                           | 5.8                         | 4.8            | 4.3            | 1.6            | 500     |
| 10              | 125      | 5.0                           | 13                          | 1.7            | 4.6            | 1.7            | 330     |
| 10              | 150      | 5.3                           | 24                          | 6.5            | 3.4            | 1.4            | 230     |
| 10              | 200      | 4.8                           | 160                         | 1.1            | 1.2            | 1.4            | 85      |
| 5               | 50       | 4.1                           | 0.04                        | 9.1            | 2.1            | 1.2            | 5,000   |
| 1               | 50       | 4.0                           | 0.005                        | 8.8            | 1.9            | 1.0            | 13,000  |

For each experiment, the post-mixing concentrations of RNA polymerase and pAR1319 DNA were 3.2 and 16 nM, respectively. The rate constants $k_{on}$, $k_{off}$, $k_m$, and $D_0$ were calculated using the mathematical model described under "Experimental Procedures." The mean number of base pairs searched per nonspecific lifetime, $I_n$, was calculated from $k_{on}$ and $D_0$ as described in the text. All experiments were performed at 37 °C.
mean number of base pairs searched per nonspecific lifetime (see Table I and "Discussion"), as evidenced by the markedly decreased total nonspecific binding at elevated salt concentrations. In Fig. 3, the value of $k_\alpha$ obtained from the curve representing binding to Fragment B is plotted as a function of KCl concentration. At low (≤50 mM) KCl concentrations, $k_\alpha$ is essentially constant at $-2 \times 10^7$ M$^{-1}$ s$^{-1}$, from 100 to 125 mM KCl, $k_\alpha$ increases approximately 2-fold to reach a maximum value, and at 200 mM KCl, promoter search slows to approximately half the rate that was measured at low salt concentrations. We were further able to calculate the nonspecific intradomain association rate constant, $k_{\alpha n}$, the intradomain dissociation rate constant, $k_{\alpha off}$, and the one-dimensional diffusion coefficient, $D_1$, utilizing the numerical methods described above. Over the range of salt concentrations studied, both $k_{\alpha n}$ and $D_1$ remain fairly constant at $-4.5 \times 10^6$ M$^{-1}$ s$^{-1}$ and $-1.5 \times 10^{-8}$ cm$^2$ s$^{-1}$, respectively. Thus the changes in the rate of promoter search observed are due almost entirely to the salt dependence of $k_{\alpha n}$. Fig. 4 shows a log-log plot of the calculated values of the equilibrium association constant for RPase binding to nonspecific DNA ($K_{eq} = k_{\alpha n}/k_{\alpha off}$) versus monovalent cation (i.e. KCl) concentration. The lines in the graph were generated using Equation 3 (17).

$$
\log K_{eq} = \log K_{eq}(1 M) - \omega \log [M']
$$

(3)

where $M'$ is the number of monovalent cations released from the DNA upon protein binding; $\omega = 0.88$ is a screening factor dependent only on the linear charge density of the DNA molecule; and $K_{eq}$ is the monovalent cation-dependent equilibrium association constant for Mg$^{2+}$ binding to DNA (16, 17). To fit the experimental data to Equation 3, log $K_{eq}(1 M)$ and $M'$ were calculated to be $-3.15$ and $12.7$, respectively. The values for the rate constants at the salt concentrations studied are summarized in Table I. A series of experiments were also performed in which NaCl was substituted for KCl, with essentially identical results (data not shown), indicating that this is a general salt effect and not due to the choice of cation.

Effect of Magnesium Ion Concentration—Since it is well known that magnesium ions profoundly weaken entropically driven DNA-protein interactions via competition with the protein for binding sites on the DNA (16, 17), we have examined its effect on the kinetics of promoter search over the range 1–10 mM. The concentration of KCl in the binding buffer was fixed at 50 mM, and all experiments were performed at 37 °C. As shown in Table I, this had a minimal effect on

![Figure 3](image-url)

**FIG. 3.** The salt dependence of the association rate constant ($k_\alpha$) of RPase to the A2 promoter-containing fragment. Curve B from each of the panels in Fig. 2, and the four other KCl concentrations examined (not shown) were fit to an exponential function (5), and the calculated rate constants are plotted against [KCl] (solid line). For all trials, [MgCl$_2$] = 10 mM, and the temperature was 37 °C. The dashed line shows the theoretical curve, calculated according to Equation 3.

![Figure 4](image-url)

**FIG. 4.** The salt dependence of the calculated intradomain, nonspecific equilibrium association constant. The rate constants $k_{\alpha n}$ and $k_{\alpha off}$ were fit to the experimental data (Fig. 2) using the mathematical model as described under "Experimental Procedures," and were used to calculate $K_{eq}(=k_{\alpha n}/k_{\alpha off})$. The two lines through the data points were generated using Equation 3; ---, the values for log$[K_{eq}]$ at 1 mM KCl and $M'=(-2.5$ and 12.3, respectively) obtained by Dehaseth et al. (18); ---, the values for log$[K_{eq}]$ at 1 mM KCl and $M' (= -3.15$ and 12.7, respectively) which best fit the data.

![Figure 5](image-url)

**FIG. 5.** Arrhenius plots of the temperature dependence of the rates of (A) specific association ($k_\alpha$) and (B) one-dimensional diffusion ($D_1$). The two rate constants were calculated as described in the text, and the best fit straight line obtained by linear regression analysis. From the slope of these lines, the activation energies for $k_\alpha$ and $D_1$ were calculated to be $2.2 \pm 0.4$ kcal/mol and $4.7 \pm 0.8$ kcal/mol, respectively.
Table II
The effect of temperature on the rate constants describing promoter search by RPase on pAR1319 DNA

| Temperature | \( k_{on} \times 10^{-7} \) | \( k_{off} \times 10^{8} \) | \( D_l \) | \( L_x \) |
|-------------|-----------------|-----------------|--------|--------|
| °C          | \( M^{-1} s^{-1} \) | \( s^{-1} \) | cm² s⁻¹ | bp     |
| 12          | 3.0             | 0.21            | 1.4 x 10⁶| 1.7    |
| 20          | 3.6             | 0.24            | 1.5 x 10⁶| 1.7    |
| 30          | 4.0             | 0.27            | 1.7 x 10⁶| 1.2    |
| 37          | 4.6             | 0.29            | 1.6 x 10⁶| 1.5    |

where \( D_l \) is the total concentration of DNA in base pairs, \( M \) is the number of sites on the DNA chain, and the contour length of the DNA fragment is \( 2L \).

This solution takes advantage of the known salt dependence of the equilibrium association constant for protein to nonspecific DNA \( (K_{eq}) \), with the plausible assumption that \( k_{on} \) is relatively insensitive to changing solution conditions \( (9) \). (The data in Table I in the present paper are consistent with this assumption.) It is the ability of this model to predict qualitatively the shape of these salt dependence curves that provides strong evidence for protein sliding during specific-site location and binding. However, exact quantitation of the one-dimensional diffusion coefficient, \( D_l \), is not possible since its effect on the value of \( k_{off} \) is so strongly dependent on \( k_{on} \), which is not measured directly. This problem is exacerbated by the fact that most experiments that do measure the off rate are in fact quantitating the dissociation of protein from DNA into the bulk solution rather than into the domain. This may serve to explain why the data points do not fit the dashed curve in Fig. 4, which was obtained by deHaseth et al. (18) using a column-elution method, but instead fit the solid line, which implies a 5-fold weaker extrapolated value of \( K_{eq} \) at 1 M KCl. From the observed difference in magnitude, it may be inferred that the protein is concentrated 5-fold within the domain with respect to the bulk, although the exact ratio should depend on the actual DNA molecule, as well as solution conditions. This interpretation is consistent with our earlier observation that rapid mixing causes the dissociation of DNA-RPase complexes, but that the bulk of the protein maintains some level of association with, and subsequently rebinds to, the DNA molecule to which it was originally bound (4, 5).

In the present study, we have demonstrated a salt dependence for the specific binding of E. coli RPase to the T7 A2 promoter similar to that described for the binding of lac repressor to its operator (6). More importantly, we have demonstrated that only \( k_{on} \) varies appreciably with the monovalent cation concentration, while \( D_l \) remains essentially constant (within experimental error). Thus, the predominant effect of monovalent cations is in determining the lifetime for protein bound nonspecifically to the DNA.

The free energy of binding of RPase to nonspecific DNA is derived almost entirely from the entropy associated with the displacement of ~12 condensed monovalent cations into solution from their ordered sites on the surface of the DNA molecule (16). From the equation for the solid line in Fig. 4, we confirm this number, within experimental error (12.7 versus 12.3). As RPase slides along the DNA template, one monovalent cation is displaced from the DNA at one end of the protein molecule while another recondenses to the site exposed by the sliding protein; this transient cation release/reassociation occurs virtually instantaneously with respect to the sliding process. Furthermore, the monovalent cation concentration in the immediate vicinity of the DNA molecule is substantial even at very low bulk ion concentrations, and varies only slightly with \( [M^+] \) (19, 20), due to the electric field induced by the anionic DNA molecule. Both these facts are consistent with our observation that \( D_l \) is unaffected by gross variations in monovalent cation concentration over the range studied (25–200 mM).

The fact that protein binding is primarily entropically driven by the release of cations from the surface of the DNA molecule has several specific consequences. First, the observed salt effects should be essentially independent of the choice of monovalent cation, as was confirmed by the duplication of the results using NaCl in place of KCl. Second, the binding energy should be a linear function of the logarithm of \([M^+]\).
rather than ionic strength. In the absence of divalent cations, which can compete with monovalent cations for binding sites along the DNA molecule (17), such a relationship has been observed for a variety of ligands including lac repressor (21), RNA polymerase (18), and several synthetic polypeptides (22). Since the association rate constant is relatively insensitive to $[M]$, the change in $K_a$ is due almost entirely to changes in $k_{on}$. The rate of specific binding is strongly dependent on $k_{on}$, in that more rapid dissociation from the DNA allows more sites to be sampled per unit time. An equivalent interpretation of this effect is that increasing $k_{on}$ (i.e., decreasing $K_a$) results in an elevated concentration of free RNA polymerase, which in turn enhances the rate at which bimolecular collisions between the protein and the promoter (or within sliding distance thereof) occur. Conversely, the mean number of base pairs searched (via sliding) per nonspecific lifetime varies inversely with the square root of $k_{off}$. Thus there are two competing factors that together determine the complex salt dependence of promoter binding by RNA polymerase (at sufficiently high concentrations such that the bimolecular collision between protein and DNA domain is not rate-limiting): how fast independent regions of the DNA are sampled ($\approx k_{on}$), and the mean number of base pairs searched per interaction ($L_s = \sqrt{D_s/k_{off}}$). By varying the concentration of monovalent cations, these two factors are differentially altered such that conflicting effects on $k_s$ result. At different concentrations, one or the other effect dominates and at some intermediate value a maximum is reached. At low salt concentrations, an increase in $k_{off}$ results in an increase in $k_s$, due to faster sampling (higher free RNA polymerase concentration), while at high salt concentrations a further increase results in a slowing of promoter search as a consequence of the shorter mean sliding distance.

As can be seen from Table I, the mean sliding distance ($L_s$) decreases almost 50-fold as $[KCl]$ is raised from 25 to 200 mM. Consequently, $L_s$ is only approximately two 40-bp compartments at the highest salt concentrations studied. In order to ensure that this approximation did not affect the analysis, the 200, 150, 125, and 100 mM data were fit with a model that broke the DNA down into 10-bp compartments. While this increased the computational demands of the analysis considerably, it was found that only the 200 mM results differed measurably in the two cases, while the 150 mM results were slightly different, and the 100 mM results were identical. Since the model using 10-bp compartments is more accurate (if there is any difference at all), the experimental data sets where a short sliding distance was observed fit using this model.

A comparison of Fig. 3 with the theoretical curves generated by Berg and von Hippel (3) for lac repressor binding shows that the low-salt “plateau” is not as pronounced for the binding of RNA polymerase to a plasmid ~5000 bp long as predicted for lac repressor. This is due to the sizable difference between RNA polymerase and lac repressor in binding affinity for nonspecific DNA (~1000-fold greater for lac repressor). Even at 25 mM KCl, a nonspecifically bound RNA polymerase molecule slides only an average of ~3000 bp/interaction, as opposed to the >10^4 bp searched by lac repressor during its nonspecific lifetime. When the nonspecific lifetime is already quite long, there is no advantage (i.e. increase in $k_s$) to sliding further than the length of the DNA, and similarly there is no disadvantage to fewer dissociation-reassociation cycles per unit time, since most protein molecules will encounter the specific site during the first nonspecific binding lifetime. The low salt plateau value of $k_s$ is therefore dependent on the length of the target DNA molecule—on average, each protein molecule binds once nonspecifically and slides one-half the contour length of the DNA molecule before encountering the specific site.

More striking than the effect of salt concentration on $k_s$ is the difference in the transient binding of RNA polymerase to nonspecific DNA represented by counts/min associated with Fragments A and C-G (Fig. 1). At salt concentrations below the maximum in Fig. 3, the nonspecific fragments display very different kinetics of RNA polymerase occupancy (Fig. 4A), while above the maximum they all behave very similarly (Fig. 4C). Due to the fact that the mean sliding distance ($L_s$) is quite long, relative to the fragment sizes, at low salt concentrations (and therefore small values of $k_{off}$), the promoter, which acts as a sink for RNA polymerase, can be “felt” by nonspecific sites located a long distance (linearly) away. The resulting family of nonspecific binding curves display markedly different kinetics, the relative positions of which correlate well with the distance between each fragment and the promoter. Those fragments located close to the promoter (Fragments C and F) lose RNA polymerase density most rapidly, while Fragment A loses RNA polymerase density most slowly. At high salt concentrations (i.e. large values of $k_{off}$), dissociation is sufficiently rapid to allow RNA polymerase to equilibrate among the nonspecific sites faster than the protein binds to the promoter, resulting in a family of curves that are virtually superimposable. Furthermore, as $K_a$ becomes weaker (due to increasing $k_{on}$), the amount of nonspecifically bound RNA polymerase decreases dramatically, as seen in a comparison of panels A-C in Fig. 2.

**Effect of Magnesium Ion Concentration on the Kinetics of Promoter Search**—The observation that the concentration of Mg$^2+$ ions has an effect on the rate of sliding of RNA polymerase along the DNA molecule may reflect structural changes in either the RNA polymerase or the DNA. While it is known that DNA undergoes gross, salt-induced structural changes such as B ↔ A and B ↔ Z transitions, it is unclear whether other, more subtle changes in structure occur gradually as the solution conditions are modified. On the other hand, it is certainly possible that Mg$^2+$-dependent conformational changes in the protein affect its ability to slide along the DNA. Further experiments are required to distinguish between these alternate explanations of the data.

The decrease in $D_s$ observed upon the reduction of [Mg$^2+$] is sufficient in and of itself to explain the degree to which magnesium ions perturbed the overall rate of promoter binding ($k_s$). Despite the fact that $k_{off}$ decreased between 2 and 3 orders of magnitude as [Mg$^2+$] was lowered from 10 to 1 mM, causing $L_s$ to increase approximately 20-fold, $k_s$ remained fairly constant. This is due to the relative insensitivity of $k_{off}$ to decreases in $k_{off}$ at 50 mM [KCl] (Fig. 3), the predominant effect of an incremental change in $k_{off}$ on $k_s$ (i.e. $dk_s/dk_{off}$) is only slightly negative as a consequence of the plateau at very small values of $k_{off}$. While it is not strictly accurate to interpret the Mg$^2+$ data using Fig. 3, in that $D_s$ does not change as a function of $[M]$, while it appears to be affected moderately by Mg$^2+$, the curve does qualitatively explain the observed lack of enhancement of $k_s$ as the magnesium concentration is lowered. The observed 30% decrease in $D_s$, as [Mg$^2+$] is decreased from 10 to 1 mM would by itself be expected to decrease $k_s$ by ~20%, due to the rate of sliding, rather than the mean sliding distance.

**The Effect of Temperature on the Kinetics of Promoter Search**—Like salt concentration, temperature has a measurable effect on the rate of promoter search, but in this case the variability in $D_s$ is primarily responsible, rather than changes in $k_{off}$. Over the range studied (4-37 °C), the rate of promoter binding varied 2-fold while the calculated values for $D_s$ varied...
approximately 4-fold. Consequently, the activation energy for $D_1$ calculated from the Arrhenius plots is approximately twice that calculated for $k_a$ (4.7 versus 2.2 kcal mol$^{-1}$). This is consistent with the theoretical derivation of Berg and von Hippel (3), in which it was predicted that $k_a$ should be proportional to $\sqrt{D_1}$ (Equation 3):

Similarly, Schranner and Richter (23) used an argument based on the “efficiency of evolution” to predict the relationship between the activation energy for sliding ($E_{\text{act}}$) and the free energy ($\Delta G$) of binding of a protein to both specific and nonspecific sites, concluding that $E_{\text{act}}$ “should” be approximately one-fourth the magnitude of the $\Delta G$ for specific binding and one-half as large as the $\Delta G$ for nonspecific binding. If one makes the approximation that the equilibrium association constant, $K_{\text{eq}}$, for the binding of $\text{RPase}$ to the $A2$ promoter is $\sim 10^{12}$ M$^{-1}$, and $K_{\text{eq}} \approx 10^5$ M$^{-1}$, then $\Delta G$ (specific) $\approx 16$ kcal/mol and $\Delta G$ (nonspecific) $\approx 8$ kcal/mol. Therefore, $E_{\text{act}}$ should be $\sim 4$ kcal/mol. From the Arrhenius plots of the temperature dependence of $D_1$, shown in Fig. 5B, $E_{\text{act}}$ was estimated to be 4.7 ± 0.8 kcal/mol, in good agreement with the value predicted by Schranner and Richter.

**Biological Consequences of Sliding**—By exploring the effects of a variety of experimental parameters on the kinetics of promoter search by *E. coli* DNA-dependent RNA polymerase, we have gained insight into the mechanism by which this enzyme (and possibly others like it) accomplishes the biologically important task of locating a specific binding site among many thousands of structurally similar nonspecific sites. The ability of nonspecific sites to sequester $\text{RPase}$ and other DNA-binding proteins is an important mechanism by which cells regulate the free intracellular concentration of these proteins, thereby modulating the rate at which such processes as transcription occur (24). Thus it may be that in order to take advantage of this equilibrium-based mechanism of the regulation of enzyme activity, the evolution of sliding was necessary to overcome the substantial kinetic barrier that nonspecific binding presents to specific-site location (in the absence of sliding). Furthermore, sliding provides a theoretical mechanism by which sites that are far apart from each other on the DNA template may interact, perhaps playing an additional regulatory role.

**REFERENCES**

1. Bujard, K., Niemann, A., Breunig, K., Roisch, U., Dresel, A., von Gabain, A., Gentz, R., Stuber, D., and Weiher, H. (1982) in *Promoters, Structure and Function* (Rodriguez, R. L., and Chamberlin, M. J., eds) pp. 121–140, Praeger, New York
2. Chamberlin, M. J., Rosenberg, S., and Kadesch, T. (1982) in *Promoters, Structure and Function* (Rodriguez, R. L., and Chamberlin, M. J., eds) pp. 34–53, Praeger, New York
3. Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981) *Biochemistry* 20, 6929–6948
4. Park, C. S., Wu, F. Y.-H., and Wu, C.-W. (1982) *J. Biol. Chem.* 257, 6950–6956
5. Singer, P. T., and Wu, C.-W. (1987) *J. Biol. Chem.* 262, 14178–14189
6. Winter, R. B., Berg, O. G., and von Hippel, P. H. (1981) *Biochemistry* 20, 6948–6960
7. Jack, W. E., Terry, B. J., and Modrich, P. (1982) *Proc. Natl. Acad. Sci. U. S. A.* 79, 4010–4014
8. Ebhrecht, H.-J., Pingoud, A., Urbanke, C., Maass, G., and Gualerzi, C. (1985) *J. Biol. Chem.* 260, 6160–6166
9. Wasylyk, B., Wasylyk, C., Augereau, P., and Chambon, P. (1983) *Cell* 32, 503–514
10. Wu, C.-W., Hillel, Z., and Park, C. S. (1983) *Anal. Biochem.* 128, 481–489
11. Schurr, J. M. (1979) *Biophys. Chem.* 9, 413–414
12. Chamberlin, M. J., Nieman, W. C., Wiggs, J., and Neff, N. (1979) *J. Biol. Chem.* 254, 10061–10069
13. Fehlberg, E. (1970) *Computing* 6, 61–65
14. Marquardt, D. W. (1963) *J. Soc. Industrial Appl. Math.* 11, 431–439
15. Record, M. T., Jr., Lohman, T. M., and deHaseth, P. L. (1976) *J. Mol. Biol.* 107, 145–158
16. Record, M. T., Jr., deHaseth, P. L., and Lohman, T. M. (1977) *Biochemistry* 16, 4790–4796
17. deHaseth, P. L., Lohman, T. M., Burgess, R. R., and Record, M. T., Jr. (1978) *Biochemistry* 17, 1612–1622
18. Manning, G. S. (1972) *Biopolymers* 11, 337–349
19. Anderson, C. F., Record, M. T., Jr., and Hart, P. A. (1978) *Biophys. Chem.* 7, 301–316
20. deHaseth, P. L., Lohman, T. M., and Record, M. T., Jr. (1977) *Biochemistry* 16, 4783–4790
21. Lohman, T. M., deHaseth, P. L., and Record, M. T., Jr. (1980) *Biochemistry* 19, 3522–3530
22. Schranner, R., and Richter, P. H. (1987) *Biochim. Biophys. Acta* 884, 135–150