Proc. Nat. Acad. Sci. USA
Vol. 69, No. 3, pp. 761–764, March 1972

Lac Repressor Binding to Synthetic DNAs of Defined Nucleotide Sequence

(nearest-neighbor analysis/rate competition experiments/equilibrium dissociation constants)

A. D. RIGGS, S. LIN, AND R. D. WELLS*

Division of Biology, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, California 91010; and
*Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wis. 53706

Communicated by Jerome Vinograd, December 27, 1971

ABSTRACT **Binding of *lac* repressor to 20 synthetic DNAs of high molecular weight with defined repeating sequences was investigated by competition experiments. Although none of these DNAs binds repressor as tightly as does *lac* operator, most do bind to a measurable extent. Their affinity for repressor varies greatly and is a function of both nucleotide composition and sequence. Poly(dC-dC)·poly(dG-dC) competes for repressor 200-times less well than either poly(dA-dT)·poly(dA-dT) or poly(dT-dT-dG)·poly(dC-dA-dA). The other DNAs show a broad spectrum of affinities for repressor between these extremes. These results show that the *lac* repressor has affinity for, and can distinguish between, sequences distantly related to its operator.**

The *lac* repressor binds tightly and specifically to the *lac* operator (1-3). However, repressor does have measurable affinity for nonoperator DNA, that is, DNA not containing the identical sequence of bases of the *lac* operator (4). In particular, repressor has a remarkably high affinity for poly(dA-dT)·poly(dA-dT): 0.01 μg/ml of this DNA will bind about 50% of the repressor present at a total concentration of 3×10^{-4} μg/ml (4). Thus, poly(dA-dT)·poly(dA-dT) provides a base sequence and/or structure recognized by *lac* repressor.

We studied other synthetic DNAs of defined sequence to gain additional information about the mechanism of this sequence-specific protein–nucleic acid interaction.

MATERIALS AND METHODS

All synthetic DNAs were prepared and characterized as described (5-7) and have molecular weights greater than 2×10^5, except for the repeating tetranucleotide DNAs that have molecular weights close to 1×10^5. Poly(dT-dT-dT-dC)·poly(dG-dA-dA-dA) has not been reported, so a nearest-neighbor analysis of this polymer is shown in Table 1. The agreement between the experimental data and expected results is good, but not perfect as found for the other DNAs used herein. *Lac* repressor and [^{32}P]λh80*dlac* DNA (molecular weight = 30×10^6; carrying *lac* operator) were prepared (3).

^{32}P-labeled repressor-operator complex was measured by filtration through nitrocellulose membranes (3). The basic experimental procedure is as follows. 10 filters were arranged on a filtration apparatus. Then, 50 μl of repressor solution (2.9 femtomol of active repressor) was added to 1.4 ml of Buffer I [10 mM KCl-3 mM Mg $(OAc)_2$-0.1 mM EDTA-0.1 mM dithiothreitol-50 μg/ml bovine serum albumin-5% (v/v) dimethyl sulfoxide-10 mM Tris·HCl (pH 7.4) at room temperature)] containing 0.082 μg of λh80*dlac* [^{32}P]DNA and the appropriate amount (0.06-9 μg/ml) of unlabeled synthetic DNA. After the solution was mixed for 5 sec with the pipette tip, 0.1-ml aliquots were filtered at 10-sec intervals. The pressure was such that filtration took about 15 sec. The filters were washed once with 0.4 ml of Buffer I without bovine serum albumin or dithiothreitol, dried, and counted. As a check on repressor stability, a control without unlabeled competing DNA was done both before and after each competition experiment. To obtain the filter-bound counts at equilibrium, additional aliquots were filtered after 10 min or more; to obtain the background counts, 0.1 mM isopropyl-β-D-thiogalactoside was added, and after 5 min three additional aliquots were filtered. The first point in the time series was assigned the value of 20 sec. Though arbitrary, this is a reasonable value that allows for mixing and filtering times. The results do not depend on the absolute time, but rather on the time intervals, which, after the first point, are accurately known. The concentration of active repressor was determined by the method of Riggs *et al.* (3). DNA concentrations were calculated from absorbance measurements and the appropriate extinction coefficients (5).

RESULTS

It has not been possible to use direct methods to measure binding of *lac* repressor to DNA that does not contain *lac* operator. However, binding of repressor to nonoperator DNA

TABLE 1. *Nearest-neighbor analysis of poly(dT-dT-dT-dC)·poly(dG-dA-dA-dA)*

[α-^{32}P] Triphosphate	% Radioactivity in dAp (a)	dAp (b)	dGp (a)	dGp (b)	dCp (a)	dCp (b)	dTp (a)	dTp (b)
dATP	63.8	60.9	36.2	30.8	0		0	8.3
dGTP	100	100	0		0		0	
dCTP	0		0	2.2	10.7	0	89.3	97.8
dTTP	0		0		41.4		58.6	

This DNA was prepared by described methods (7). A mixture of $(dT\text{-}dT\text{-}dT\text{-}dC)_2$ + $(dA\text{-}dA\text{-}dA\text{-}dG)_3$ (8) was used as template (primer) in one set of experiments (*a*); a mixture of $(dT\text{-}dT\text{-}dT\text{-}dC)_3$ + $(dA\text{-}dA\text{-}dA\text{-}dG)_4$ was the template (primer) in another series of experiments (*b*); experiment not done with [α-^{32}P] dTTP). Nearest-neighbor analyses were performed (9); more than 10,000 cpm of labeled polymer was degraded for each determination.

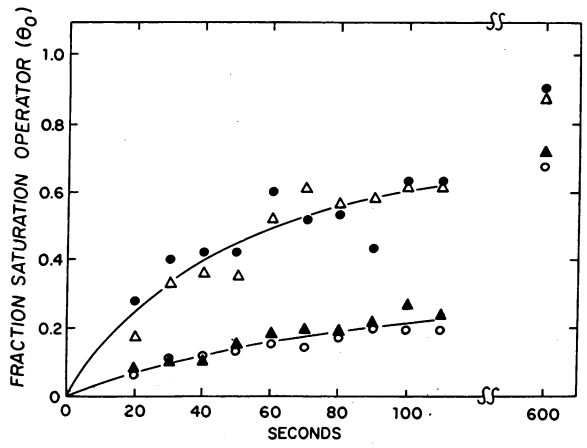

FIG. 1. Kinetics of formation of repressor–operator complex in the presence of competing DNA. The experimental procedure is described in *Methods*. ●——●, control, no competing DNA; △——△, 0.12 μg/ml poly(dI-dC)·poly(dI-dC); ○——○, 0.12 μg/ml poly(dA-dT)·poly(dA-dT); ▲——▲, 0.12 μg/ml poly-(dT-dT-dG)·poly(dC-dA-dA).

can be measured by competition experiments. As conventionally performed, the effect of unlabeled competing DNA on the equilibrium concentration of [^{32}P]repressor–operator complex is measured. Many detailed equilibrium competition experiments have, in fact, been done on repressor binding to nonoperator DNAs from natural sources and to synthetic poly(dA-dT)·poly(dA-dT) (4, Lin and Riggs, unpublished data).

Because of the extremely limited supply of fully characterized synthetic DNAs of defined repeating sequence, we did not do equilibrium competition experiments. Instead, more sensitive, but somewhat unconventional, *rate competion experiments* were used. The rate of formation of [^{32}P]repressor–operator complex can be conveniently and accurately measured, and it was found that second-order kinetics are obeyed (10). Competing DNA that binds repressor and reduces the concentration of free repressor would be expected to decrease the rate of formation of [^{32}P]repressor-operator complex. Fig. 1 shows that this is indeed the case. Both poly(dA-dT)·poly(dA-dT) and poly(dT-dT-dG)·poly(dC-dA-dA) at 0.12 μg/ml greatly decrease the rate at which the complex is formed, whereas poly(dI-dC)·poly(dI-dC) at the same concentration has no effect. Rate competition experiments are about ten times more sensitive than equilibrium competition experiments. At these concentrations, poly(dA-dT)·poly-(dA-dT), for example, only slightly affects the equilibrium concentration of the repressor-operator complex (Fig. 1, 10-min point).

For each DNA, it is desirable to make quantitative estimates of K_{RD}, the equilibrium dissociation constant that characterizes binding of nonoperator DNA. Therefore, the appropriate theoretical treatment for rate competition experiments was developed (Lin, and Riggs, unpublished). We will present a brief outline of the method here.

Because of the low concentrations used, binding of repressor to operator is very slow and takes many seconds. Therefore, we assume that binding of repressor to competing nonoperator DNA is essentially at equilibrium throughout the rate competition experiment. This being the case, it can be shown that the presence of competing nonoperator DNA changes the apparent second order rate constant for repressor–operator association, i.e.,

$$dRO/dt = k_a \left(\frac{K_{RD}}{K_{RD} + D_t}\right)(R_f)(O_f), \qquad [1]$$

where RO is the repressor–operator complex, R_f and O_f are free repressor and free operator, respectively, and k_a is the bimolecular rate constant for repressor–operator association. D_t is the total number of nonoperator binding sites. Because of the resulting automatic correction for filter retention efficiency and changing specific activity of ^{32}P, we have found it convenient to use θ_0, the fractional saturation of operator by repressor, as our basic parameter.

$$\theta_0 = RO/O_t = cpm/cpm_{max}, \qquad [2]$$

where cpm_{max} is the filter-bound counts obtained when operator is fully saturated with repressor. Integrating Eq. (1) for the case where O_f equals R_f, then substituting θ_0 for RO, one can obtain

$$1/(1 - \theta_0) = k_a \left(\frac{K_{RD}}{K_{RD} + D_t}\right) O_t t + 1 \qquad [3]$$

In agreement with Eq. (3), straight lines are obtained when the data from Fig. 1 are plotted as in Fig. 2. From the slope of the line in the presence of competing DNA and the slope of the control, K_{RD} can be determined. For poly(dA-dT)·poly(dA-dT), K_{RD} estimates obtained by this method have been carefully compared with those obtained by conventional equilibrium competition experiments; good agreement was found.

Experiments such as those shown in Figs. 1 and 2 were done for 20 different synthetic DNAs of defined sequence, and the outlined procedure was used to estimate K_{RD} values. Fig. 3 summarizes the results. There are clearly large differences in competition effectiveness among the synthetic DNAs tested. It should be noted that although we have used Equation (3) to obtain K_{RD} values, the basic fact that DNAs vary greatly in their ability to reduce the rate of formation of repressor–operator complex does not depend on any theoretical treatment. For example, 0.1 μg/ml of poly(dA-dT)·poly-(dA-dT) has a pronounced effect, whereas 94 times as much poly(dG-dC)·poly(dG-dC) (9.4 μg/ml) is required to have a significant effect on the rate of complex formation.

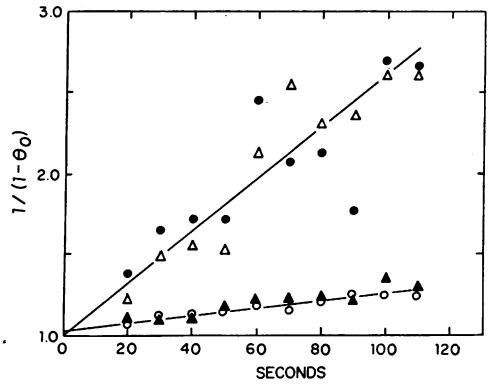

FIG. 2. Linear plot of rate competition data. The data shown in Fig. 1 were plotted so as to be linear for a second-order reaction (see *Results*). The symbols are the same as in Fig. 1.

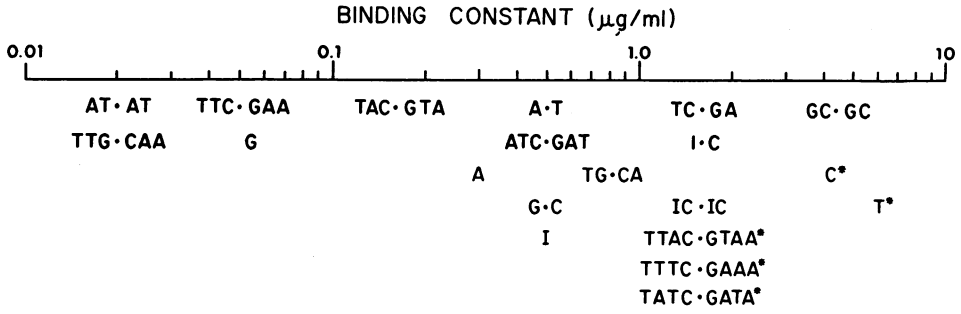

FIG. 3. Binding constants of synthetic DNAs to repressor. For each DNA, K_{RD} was estimated in units of μg/ml (see *Results* for procedure). The base sequence for each DNA is indicated and centered about the K_{RD} for that DNA (e.g., TTG-CAA is an abbreviation of the standard IUPAC nomenclature of poly(dT-dT-dG)·poly(dC-dA-dA) for this DNA; K_{RD} is 0.02 μg/ml). Every DNA was tested at least twice, usually at different concentrations. Reproducibility was good; in all cases the range of K_{RD} was less than a factor of two. For the controls, the standard deviation in the slope was ±25% for 34 Exps. The DNAs marked with (*) had no significant effect even at the highest concentration tested, which was 1.2 μg/ml for poly(dC) and poly(dT), and 0.47 μg/ml for the repeating tetramer DNAs.

Of the single-stranded DNAs, poly(dC) and poly(dT) do not compete, poly(dI) and poly(dA) compete moderately, and poly(dG) competes rather well. Perhaps repressor binds to poly(dG). However, it is also possible that these single-stranded DNAs interact with operator DNA instead of with repressor and, hence, give anomalous results. This latter notion is supported by the observation that the single-stranded polymers compete in roughly the same order as their ability to aggregate with themselves or other DNAs (refs 5, 11, 12; Wells, unpublished work). Experiments to test these two possibilities have not been definitive. Thus, the significance of their competition effectiveness can not be ascertained at present. Because of the interstrand complementarity and double-helical character (5) of the other DNAs tested, it is extremely unlikely that they are interacting with operator, which is double-stranded (3), rather than with repressor.

DISCUSSION

These studies demonstrate that some DNA polymers effectively compete against operator DNA in binding to the *lac* repressor. A 200-fold range in the apparent affinity of synthetic double-stranded DNAs for the repressor was found. Poly(dG-dC)·poly(dG-dC) was the poorest competitor (K_{RD} = 4 μg/ml), whereas poly(dA-dT)·poly(dA-dT) and poly(dT-dT-dG)·poly(dC-dA-dA) were the best competitors (K_{RD} = 0.02 μg/ml for each).

An unambiguous assessment of the features necessary for the binding of a DNA to the *lac* repressor is difficult. We have previously established that certain primary nucleotide sequences affect the secondary structure of a DNA (5, 13). Since we poorly understand the manner by which this occurs, and since a wide spectrum of binding capabilities was found for the DNAs (Fig. 3), it is difficult to ascertain precise structural features that are responsible for the binding. Only DNAs rich in (dA + dT) bind well to repressor; however, not all DNAs rich in (dA + dT) are effective competitors. Both poly(dA-dT)·poly(dA-dT) and poly(dA)·poly(dT) have identical base composition, yet the former competes 25 times more effectively than the latter. Poly(dA-dT)·poly(dA-dT) has a fiber diffraction pattern identical to the B form of DNA (14), but poly(dA)·poly(dT) has a somewhat different structure (13). We previously considered (4) the possibility that in the alternating dA and dT polymer, repressor bound to loops or branches that exist because of the intrastrand complementarity. Our studies render this possibility unlikely since (*a*) poly(dT-dT-dG)·poly(dC-dA-dA), which does not have a self-complementary structure, binds repressor well on a weight basis and (*b*) poly(dI-dC)·poly(dI-dC) and poly(dG-dC)·poly(dG-dC), which also contain intrastrand complementarity (15), bind quite poorly.

It is not sufficient for binding to repressor that a polymer have properties similar to natural DNA. Poly(dT-dG)·poly(dC-dA) has properties quite similar to natural DNA containing 50% (dG+dC); however, it binds repressor 40 times less well than some other DNA polymers.

The four repeating trinucleotide DNAs, all containing 33% (dG + dC), have different capacities to bind repressor, ranging from K_{RD} = 0.02 μg/ml for poly(dT-dT-dG)·poly(dC-dA-dA) to K_{RD} = 0.5 μg/ml for poly(dA-dT-dC)·poly(dG-dA-dT). All four polymers have very similar properties such as melting temperatures, buoyant density values, and circular dichroism spectra (5), yet they show quite different binding abilities. Poly(dT-dT-dG)·poly(dC-dA-dA) binds much better than expected from the results with polymers that have sequences related to DNA; poly(dA)·poly(dT), and poly(dT-dG)·poly(dC-dA), are rather poor binders. That poly(dA-dT-dC)·poly(dG-dA-dT) binds less well than the other three repeating trinucleotide DNAs is expected since studies of actinomycin D binding (12) indicated that it possesses unique properties. There is no direct correlation between the melting points of the DNAs (5) and their affinity for repressor.

It was unexpected that the repeating tetranucleotide DNAs showed no binding. The reason for this is unclear. It may be related to their low molecular weight (7) relative to the other polymers, but it is more likely due to other considerations (see below) because they have molecular weights of 1×10^5. A DNA of this size contains 160 base-pairs, and the *lac* operator is not likely to be greater than 20 base-pairs.

In considering the total binding pattern observed, we believe that more than structural differences are involved. Apparently, repressor is interacting with the nucleotides in a sequence-specific manner.

None of the DNAs studied bind repressor nearly as tightly as does the *lac* operator. When synthetic DNAs are used, the molar concentration of binding sites is always much greater than the concentration of repressor. Therefore, to estimate

K_{RD} in molar terms one must consider the concentration of all potential binding sites. For poly(dA-dT)·poly(dA-dT), every base-pair can start a binding site, and thus the effective molar concentration of binding sites for this DNA equals the concentration of base-pairs. From these experiments, and others not shown, we estimate that K_{RD} for poly(dA-dT)·poly(dA-dT) is 10–30 nM. For poly(dT-dT-dG)·poly(dC-dA-dA), only every third base-pair begins a binding site, so K_{RD} is estimated to be 3–10 nM. In absolute terms, this very tight binding, but under similar conditions the equilibrium dissociation constant for repressor-operator binding is $2\text{–}5 \times 10^{-5}$ nM (11).

Poly(dA-dT)·poly(dA-dT) and poly(dT-dT-dG)·poly(dC-dA-dA) do not seem at all related in terms of sequence. However, they may resemble sequences occurring at different regions of the operator. It seems clear that the *lac* repressor has partial affinity for partially correct sequences. One simple model that would be expected to give this result is that the total DNA binding site of the *lac* repressor is composed of several subsites or binding elements, each binding for the most part independently to its preferred short sequence of bases. It may be significant that repeating tetramers are not effective binders (Fig. 3). Poly(dT-dA-dC)·poly(dG-dT-dA) binds repressor with $K_{RD} = 0.16$ μg/ml. Addition of one extra thymidine to give poly(dT-dT-dA-dC)·poly(dG-dT-dA-dA) decreases repressor binding more than 10-fold. Perhaps the repeating (dT-dA-dC) sequence is shifted out of phase.

These studies are significant with regard to mechanisms of protein–nucleic acid interactions and to the evolution of operators and regulatory systems. Since partially correct sequences have partial affinity for repressor, pseudo-operator sites must occur with reasonable frequency in natural DNAs, and this suggestion probably explains why repressor does have some affinity for most natural DNAs (4). By binding repressor, nonoperator DNA can greatly reduce the apparent affinity of repressor for operator; such binding may have important physiologic consequences.

This work was supported by grants from the National Institutes of Health (1-R01-HD-04420) and the National Science Foundation (GB-26517) to A. D. R., and by grants from the National Science Foundation (GB-8786) and the Damon Runyon Memorial Fund (DRG-1049) to R. D. W.

1. Gilbert, W. & Müller-Hill, B. (1967) *Proc. Nat. Acad. Sci. USA* **58**, 2415–2421.
2. Riggs, A. D., Bourgeois, S., Newby, R. F. & Cohn, M. (1968) *J. Mol. Biol.* **34**, 365–368.
3. Riggs, A. D., Suzuki, H. & Bourgeois, S. (1970) *J. Mol. Biol.* **48**, 67–83.
4. Lin, S. & Riggs, A. D. (1970) *Nature* **228**, 1184–1186.
5. Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E. & Cantor, C. R., (1970) *J. Mol. Biol.* **54**, 465–497.
6. Burd, J. F. & Wells, R. D. (1970) *J. Mol. Biol.* **53**, 435–460.
7. Wells, R. D., Buchi, H., Kossel, H., Ohtsuka, E. & Khorana, H. G. (1967) *J. Mol. Biol.* **27**, 265–272.
8. Jacob, T. M., Narang, S. A. & Khorana, H. G. (1967) *J. Amer. Chem. Soc.* **89**, 2177–2184.
9. Wells, R. D., Jacob, T. M., Narang, S. A. & Khorana, H. G. (1967) *J. Mol. Biol.* **27**, 237–263.
10. Riggs, A. D., Bourgeois, S. & Cohn M. (1970) *J. Mol. Biol.* **53**, 401–418.
11. Inman, R. B. (1964) *J. Mol. Biol.* **9**, 624–637.
12. Wells, R. D. & Larson, J. E. (1970) *J. Mol. Biol.* **49**, 319–342.
13. Langridge, R. (1969) *J. Cell. Physiol.* **74**, Suppl. 1, 1–20.
14. Davies, D. R. & Baldwin, R. L. (1963) *J. Mol. Biol.* **6**, 251–255.
15. Grant, R. C., Kodama, M. & Wells, R. D. (1972) *Biochemistry*, in press.