INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392-1

SMR/111 - 16

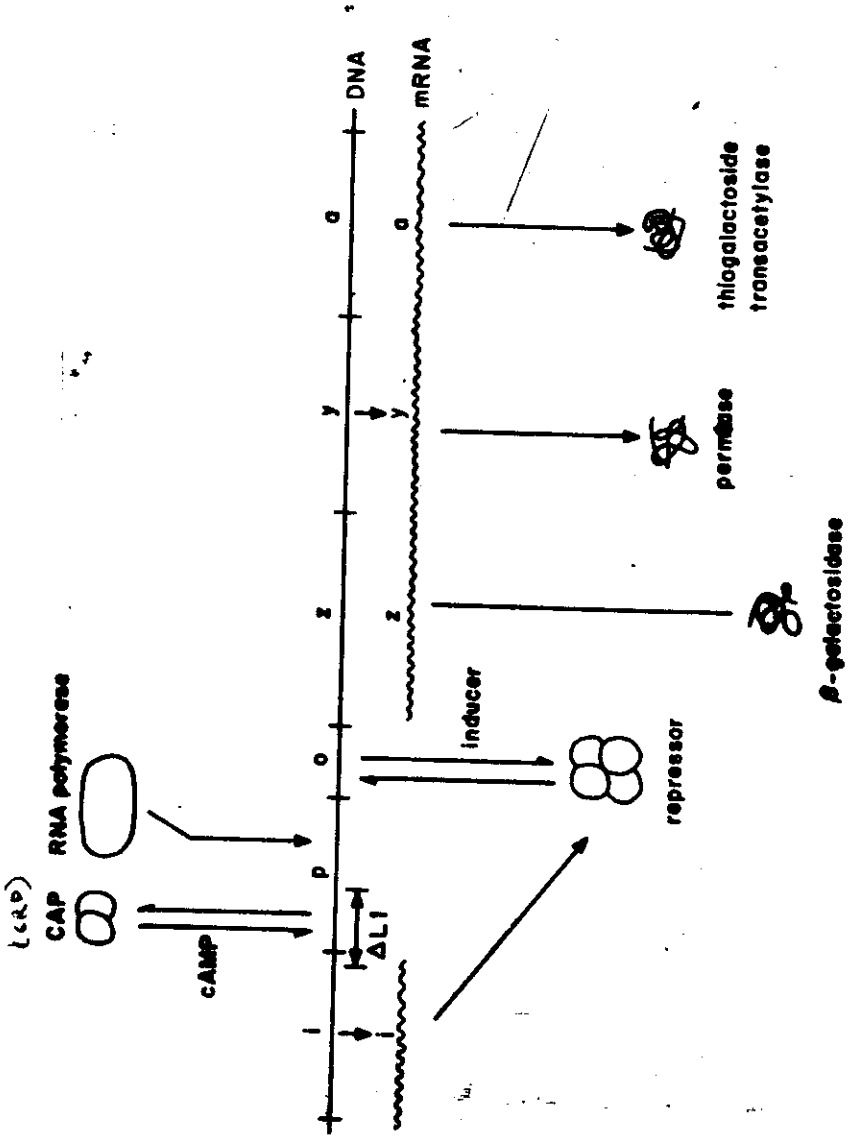
SECOND SUMMER COLLEGE IN BIOPHYSICS

30 July - 7 September 1984

MISC. TRANSPARENCIES

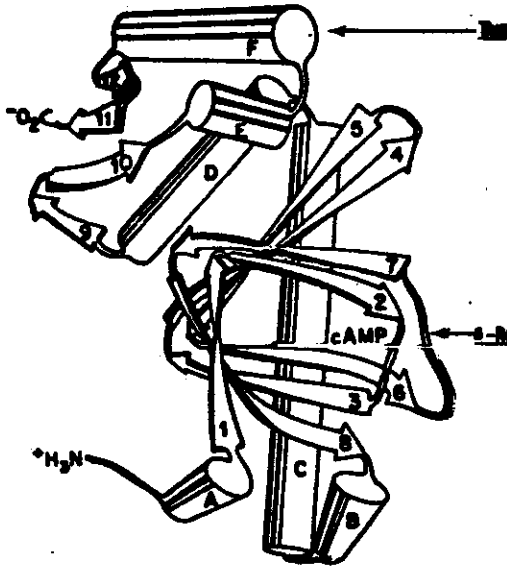
D.M. CROTHERS

Department of Chemistry
Yale University
225 Prospect Street
P.O. Box 6666
New Haven, Connecticut 06511
U.S.A.



Weber and Steitz
Figure 1

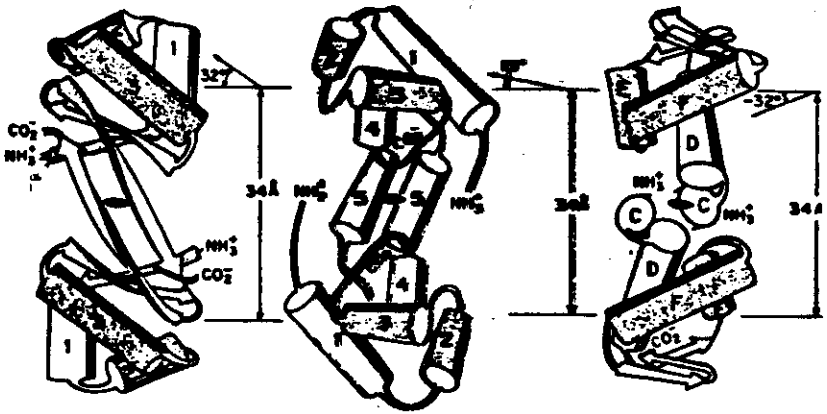
CAP HAS A MODULAR STRUCTURE



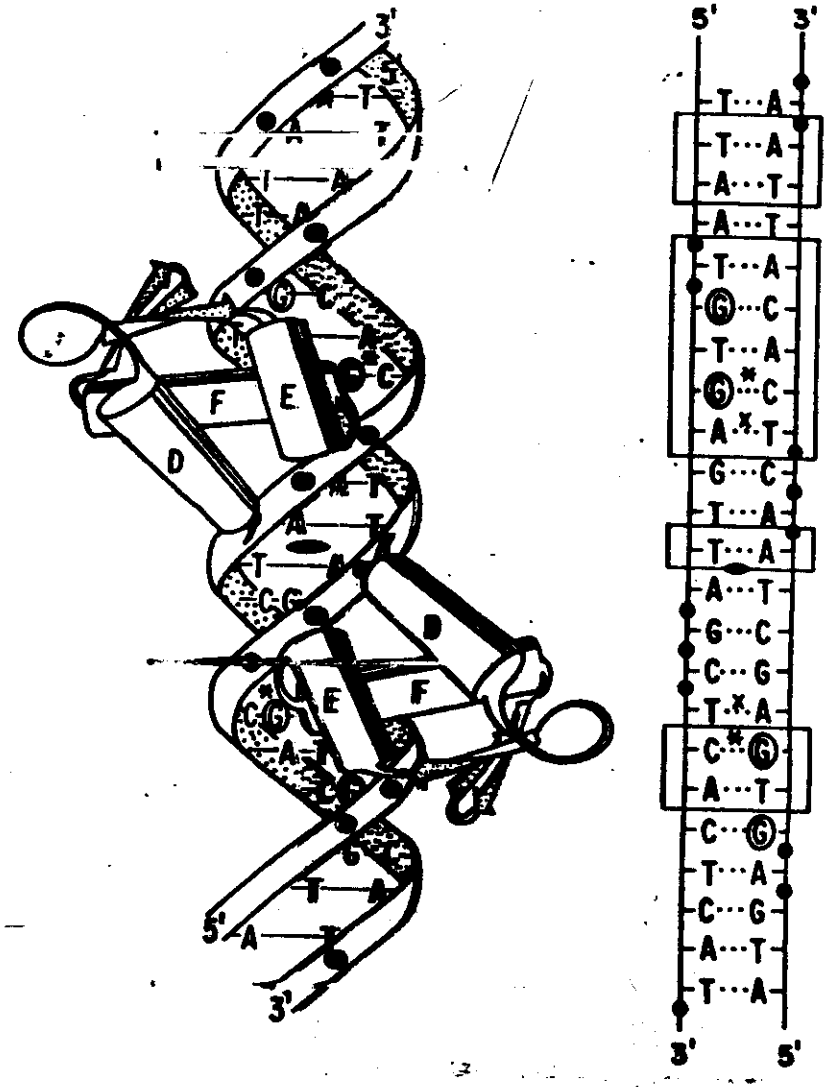
HTH-HELIX MOTIF
STRUCTURE IDENTITY
λ CRO
λ CI
SEQUENCE HOMOLOGY
LAC REPRESSOR
GAL REPRESSOR
RESOLVASE
OTHER REPRESSORS

α-ROLL STRUCTURE
STRUCTURE ANALOGY
VIRUS COAT PROTEINS
HEMAGGLUTININ
SEQUENCE HOMOLOGY
REGULATORY SUBUNIT
cAMP DEPENDENT
PROTEIN KINASE

Weber and Steitz
Figure 8



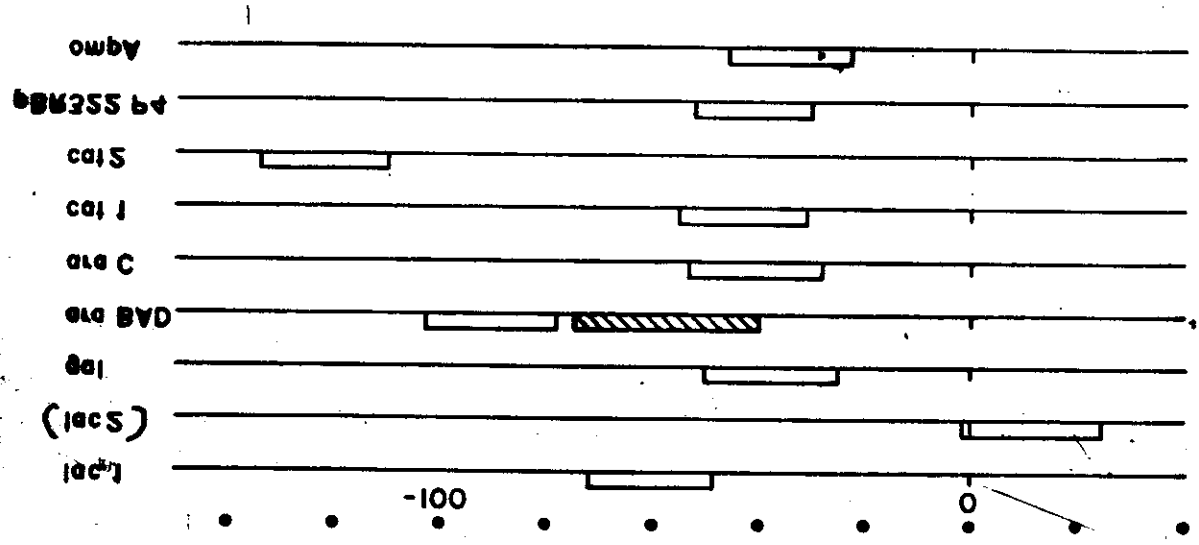
Weber and Steitz
Figure 11



PROMOTER

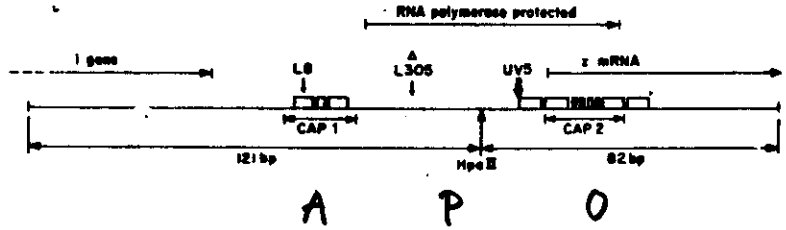
lac 1	C A A T T A A T G T G A G T T A G C T C A C T C A T T A G G
(lac 2)	(T G T G G A A T T G T G A G C G G A T A A C A A T T T C A C A)
gal	C C A C T A A T T T A T T C C A T G T C A C A C T T T T C G C A
ara BAD	T T G A T T A T T T G C A C G G C G T C A C A C T T T G C T A
ara C	C T T T T C T G C C G T G A T T A T A G A C A C T T T T G T T A
cat 1	C G A A A A T G A G A C G T T G A T C G G C A C G T A A G A G
cat 2	T A A A T A C C T G T G A C G G A A G A T C A C T T C G C A G
pBR 322 p4	A C G C A T C T G T G C G G T A T T T C A C A C C G C A T A T
amp A	T C A T A T G C C T G A C G G A G T T C A C A C T T G T A A G

CONSENSUS T G T G A - $N_{6 \pm 1}$ - T C A C A C T T T



Methods for Studying Protein-DNA Interactions (in solution)

1. Filter binding: ^{32}P-labeled DNA binds to a membrane filter only if there is protein bound to the DNA
   - highly sensitive
   - but cannot easily distinguish multiple reaction products
2. Spectroscopy
   - requires relatively large amounts of material
   - often the binding constant is too large for determination this way
   - but structural information may result, especially from NMR
3. "Footprinting": Partial enzymatic or chemical modification of the DNA portion with and without protein
   - excellent method for determining binding site
4. Gel electrophoresis of protein-DNA complexes
   - gives equilibrium & kinetic assay
   - different complexes are resolved
   - yields conformational information



UV5 —> strong promoter

Determining stoichiometry from equilibrium constant measurements

$$CAP + P + \underline{n}\text{-}cAMP \rightleftharpoons CAP\text{-}P\text{-}cAMP_n$$

$$K = \frac{[CAP\text{-}P\text{-}cAMP_n]}{[CAP][P][cAMP]^n}$$

$$= \underbrace{\frac{[CAP\text{-}P\text{-}cAMP_n]}{[CAP][P]}}_{K_{ap}} \cdot \frac{1}{[cAMP]^n}$$

$$\ln K = \ln K_{ap} - n \ln [cAMP]$$

$$\ln K_{ap} = \ln K + \underline{n} \ln [cAMP]$$

↑ measure ↑ constant

Diffusion-Limited Reactions

Reaction occurs whenever particles diffuse within a distance r_{12} of each other

(irrespective of mutual orientation). Then

$$k = \frac{4\pi N_A}{1000}(D_1 + D_2) r_{12} \cdot \alpha$$

D_1, D_2 are the diffusion constants

Electrostatic acceleration:

$$k = \frac{4\pi N_A r_{12}(D_1 + D_2)}{1000} \cdot \frac{\phi_{12}}{e^{\phi_{12}} - 1}$$

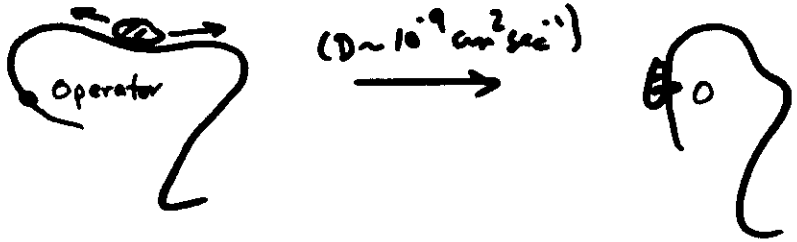
ϕ_{12} = ratio of electrostatic energy to thermal energy.

DNA - Protein Search Kinetics

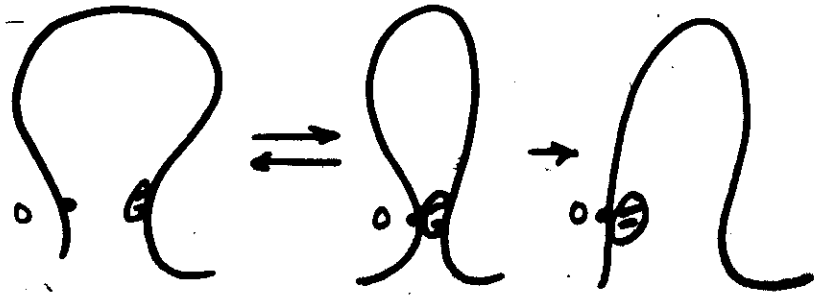
- Reaction bimolecular rate constant ($\sim 10^{10}$ M^{-1} sec^{-1}) faster than diffusion limited value ($\sim 10^{8}$ M^{-1} sec^{-1})

- General model: intramolecular search after molecular capture, by:

a) Sliding



b) Direct Transfer



proposed mechanism is called intersegment transfer[4,10], and involves the rapid and direct transfer of repressor from one segment of a DNA molecule to another as a consequence of the relative diffusion of these segments within the 'domain' of the molecule. This second mechanism postulates the transfer of repressor by a series of 'ring-closure' events in which the repressor is transiently bound between two DNA segments (current views of repressor structure suggest that the repressor tetramer may have at least two DNA binding sites[3]). When the segments diffuse apart one of the DNA–protein contacts will break and, if the two binding interactions are equally tight, the repressor will have a 50% chance of being transferred to a new and distant site on the DNA molecule. This process may be very fast because it circumvents the large activation barrier involved in the dissociation of repressor into solution.

Both mechanisms are plausible, given the known properties of *lac* repressor (see below), and both can, in principle, lead to the desired result. Fortunately, these mechanisms are experimentally distinguishable because sliding is correlated with position *along* the DNA molecule, while intersegment transfer is not.

Experiments and analysis

The binding of repressor to operator-containing DNA was examined using a filter-binding technique[11]. This method takes advantage of the fact that protein binds tightly to nitrocellulose filters, while free DNA passes through. Radioactively-labelled DNA will therefore only be retained on the filter if it is tightly complexed to protein. Conditions are fixed so that weak complexes, such as those due to non-specific binding (RD interactions), are *not* retained, while RO complex formation holds the labeled DNA on the filter. Both equilibrium and kinetic measurements can be made this way; because of their bimolecular character even very fast association reactions can be brought into the experimental time range by dilution. Careful kinetic and equilibrium measurements of the RO interaction have been made by this technique[4,7,10]. We focus here on the association process.

In Fig. 2 we show some typical values of k_+ (plotted as log k_+), as a function of KCl concentration (plotted as log [KCl] or log K_{RD}), for three different sizes of *lac* operator-containing DNA fragments. The resulting values of k_+ depend in a unique manner on DNA length and salt concentra-

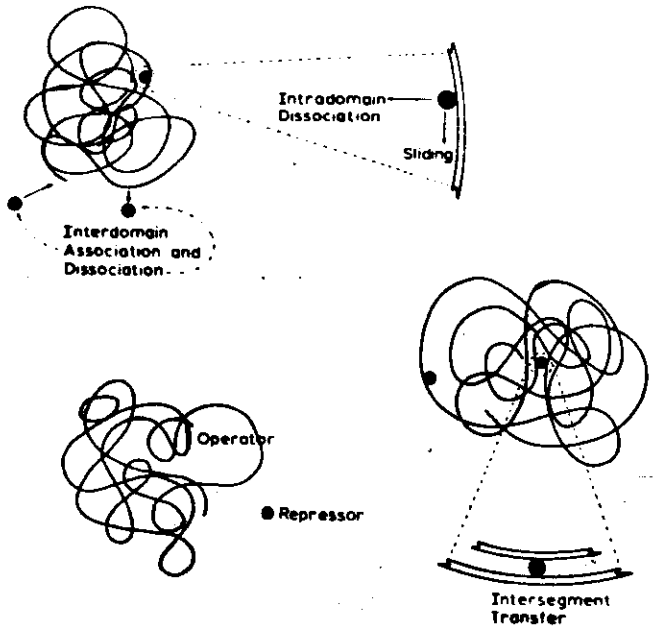


Fig. 1. Schematic view of lac *repressor (R) interacting with large operator-containing (O) DNA molecules in dilute solution. The DNA molecules are well-separated into 'domains' under these conditions. The (upper right) expanded view shows repressor bound to a segment of non-operator DNA, on which it can either 'slide' or engage in interdomain dissociation–association processes in seeking its specific (operator) target site. The (lower right) expanded view shows a repressor molecule doubly bound to two DNA segments; this corresponds to the (hypothetical) intermediate state in the intersegment transfer process.*

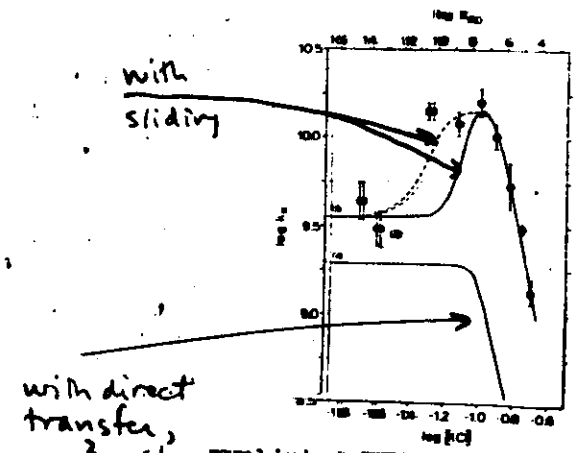


FIGURE 2: Log k_a vs. log [KCl] for RO complex formation with λplac5 DNA (see Materials and Methods). The points represent the experimental data, with narrow bars indicating the standard deviation for experiments repeated under identical conditions. The lines are theoretical and have been calculated as described in Berg et al. (1981) and in the text. The solid curves represent the predicted RO association constants calculated as a function of N_{ion} (upper abscissa) estimated by using eq 13. The parameter values used are $D = 5 \times 10^{-7}$ cm²/s, $M = 5 \times 10^4$ base pairs, R_c = 1200 Å, r_0 = 5500 Å, and $O_T = 10^{-12}$ M. Curve a shows the results expected using the intersegment transfer mechanism as the sole translocation mechanism, calculated according to eq 7 and 11 with $\gamma = 100\ s^{-1}$. Curve b shows the results expected using the sliding mechanism only, according to eq 7 and 12 with $D_1 = 9 \times 10^{-10}$ cm²/s. The dashed portion of curve b is calculated assuming a weaker salt dependence of K_{RO} at low salt (i.e., $\log K_{RO} = -5 \log [KCl] + 2.25$ at $\log [KCl] < -0.95$; see also footnote 2).

(Winter, Berg, & von Hippel)
(Biochem 20, 6961 [1981])

Intramolecular Direct Transfer

$$R\text{-}D + O \underset{k_R}{\overset{k_F}{\rightleftharpoons}} R\text{-}O + D$$

(R = repressor; O = operator; D = non-specific DNA)

$$K = \frac{k_F}{k_R} = 3 \times 10^6\ M^{-1}$$ (Lin & Riggs, 1975)

$k_R = 30\ M^{-1}\ sec^{-1}$ (Fried, 1982)

$\therefore\ k_F = 10^8\ M^{-1}\ sec^{-1}$ (agrees with results of Winter et al, 1981)

Intramolecular transfer rate = k_f × local site conc.

(site conc. $\sim 10^{-4}$ M from data of Shore Langowski & Baldwin, 1981)

Therefore intramolecular rate = $10^4\ sec^{-1}$

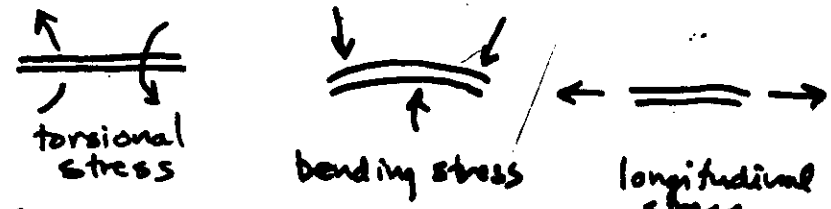
Average time between transfer events $\sim 10^{-4}$ sec

Sliding distance between transfer events 10-30 bp

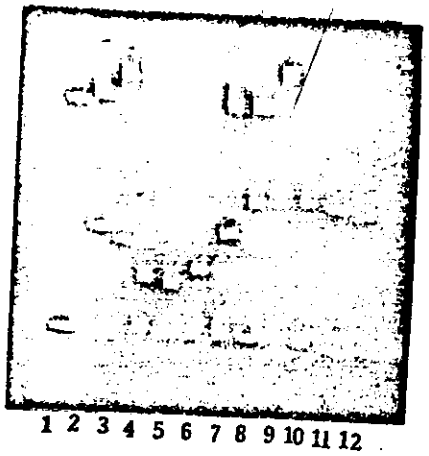
Assembly – Disassembly

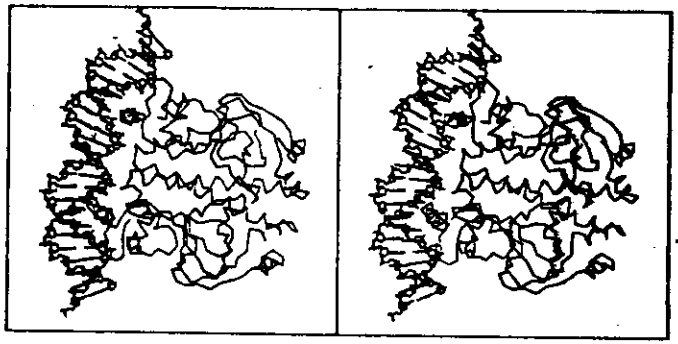
(Participation of distal DNA sequences in formation of specific protein-DNA complex)

Disassembly → ← Assembly

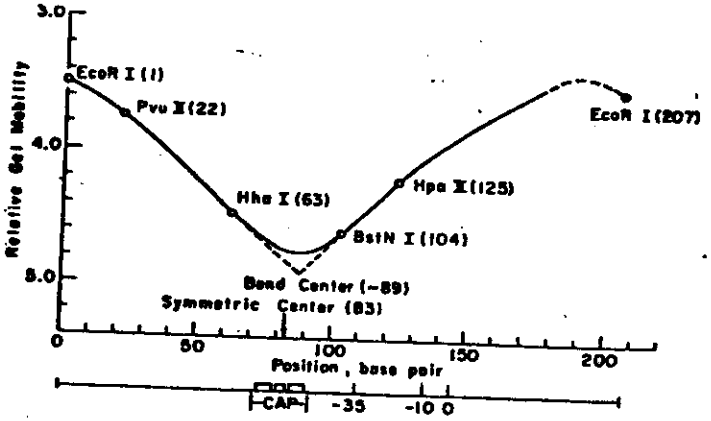


a



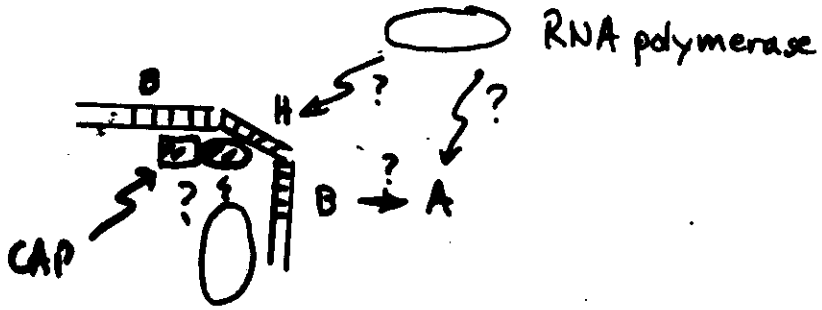


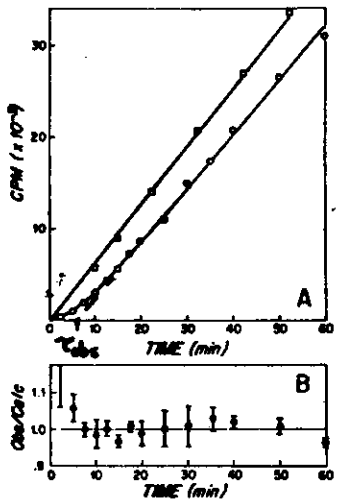
b



Weber and Steitz
Figure 12

Possible structural models





$$R + P \underset{k_{-B}}{\overset{K_B}{\rightleftharpoons}} RP \underset{k_{-c}}{\overset{k_c}{\rightleftharpoons}} RP_o$$

$$\tau_{obs} = \frac{1}{k_c} + \frac{1}{K_B[R]k_c}$$

FIG. 3.1. The time required for open-complex formation on the prmup-1 promoter. The standard assay conditions described in Table 3.1 were employed. A. Radioactivity incorporated into UpApU is plotted versus time. The reaction corresponding to the linear time course was initiated with UpA and [α-^{32}P]UTP following preincubation of RNA polymerase and the promoter-containing DNA fragment (□). The reaction initiated with RNA polymerase (25 nM) showed a lag (τ_{obs} = 6 min) before the final steady state reaction rate was achieved (○). The curve was calculated with a nonlinear, least squares minimization program. B. The lag experiment of panel A was performed at four RNA polymerase concentrations (i.e., to generate part of a TAU plot). For each curve, the observed incorporation values were divided by the calculated values at each time point. The distribution of these ratios around the ideal value of 1.00 provides an estimate of the relative error in the experiments. At each time point, the mean ± S.D. is indicated.

Hawley, Nalan, Mulligan & McClure

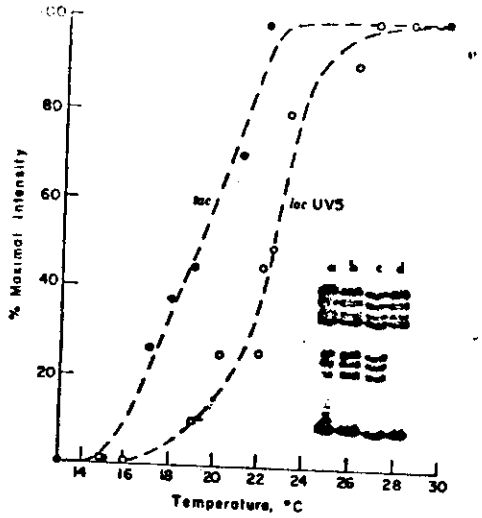


FIG. 4. Temperature dependence of cytosine methylation in the presence of RNA polymerase for *lac* UV5 and *tac* promoters. The extent of methylation of the cytosines in the −10 region of the *tac* (●) and *lac* UV5 (○) promoters by dimethyl sulfate in the presence of bound RNA polymerase was measured as a function of incubation temperature. At each temperature, saturation of the promoter by bound RNA polymerase was verified by DNase I protection. The length of incubation of RNA polymerase–promoter complexes with 200 mM dimethyl sulfate was varied from 5 min at 30°C to 20 min at 10°C in order to achieve approximately the same extent of methylation for each sample. Hydrazine and piperidine treatments and sequencing gel electrophoresis were performed identically for each sample. The resulting autoradiograms were scanned with a densitometer. The extent of cytosine methylation in each sample was determined by normalizing the average band intensities from the four cytosines at positions −1, −2, −4, and −6 to the band intensities from guanine residues represented by fragments both longer and shorter on the sequencing gel (guanines at +17, −16, −19, or −20). Values for the extent of cytosine methylation in the presence of RNA polymerase below 15°C were indistinguishable from controls in the absence of the enzyme for both promoters: this value was defined as 0% maximal intensity on the graph. For each promoter, there was a temperature beyond which the calculated extent of cytosine methylation no longer increased, the value at this point was defined as 100% maximal intensity. Intermediate values were normalized to a linear scale within this 0–100% range; the error for each point is estimated to be about 10%. For the *lac* UV5 promoter, this curve is unaffected by incubating the RNA polymerase–promoter complex at 37°C before incubating at the given temperature. (*Inset*) Portion of a sequencing gel displaying some of the data from which this graph was derived. A 400-base-pair fragment bearing the *tac* promoter (19, 20), uniquely end-labeled at the 3′ end of the template strand, was incubated at 18°C (lane a), 22°C (lane b), and 30°C (lanes c and d) in the absence (lane d) and presence (lanes a, b, and c) of saturating amounts of RNA polymerase. Incubations were continued in the presence of 200 mM dimethyl sulfate for 6, 4, and 2 min for the 18°C, 22°C, and 30°C reactions.

polymerase complex corresponding to the guanine residue at position −4 of the *lac* UV5 promoter is notably better protected from dimethyl sulfate, suggesting the presence of an additional promoter–polymerase contact in the *tac* complex that is absent from the *lac* UV5 complex.

For the *trp* promoter, although several cytosines are present in the region of positions +1, −2, −6, and −11 in the template or bottom strand and at positions −2 and −9 in the top strand), we could not detect methylation of any of these cytosines at temperatures as high as 37°C. The trivial interpretation that the *trp* promoter does not form a unique complex with RNA polymerase in our experiments is made highly unlikely by the characteristic protection and enhancement patterns of the guanines (Fig. 3 and ref. 22) and by DNase I protection (data not shown). Indeed, all polymerase contacts revealed by the guanine methylation and DNase I cleavage patterns of the *tac* and *trp* promoters are identical in the −35 region, which is reasonable in view of the sequence identity of the two promoters in this region.

DISCUSSION

The results presented above demonstrate the applicability of the cytosine N-3 methylation reaction in probing the single-stranded regions in a duplex DNA. The signal-to-noise ratio of this method for detecting unpaired cytosines is limited by two major factors. The reactivity of unmethylated cytosine with hydrazine is appreciable; the resulting chain cleavages at the positions of the unmethylated cytosines therefore give a significant level of background. Also, the principle of sequencing an end-labeled linear polymer by chain breakage requires the introduction of no more than a few breaks within a distance from the labeled end. This distance is determined by the resolving power of the sizing method, typically a few hundred nucleotides for nucleic acids. Because the guanines are readily methylated, and chain breakage by piperidine occurs at the positions of the N-7-methylated guanines on a sequencing gel the maximal intensity of the bands resulting from methylation of the cytosines is about one-third of the average intensity of the bands resulting from the methylation of the guanines (see Fig. 1). These limitations on the signal and the noise tend to narrow the window for optimization of the reaction conditions; in cases in which the cytosines are unpaired only a small fraction of the time, the method is unlikely to give an unequivocal answer.

Nevertheless, the mildness of the methylation reaction and the utility of the cytosine methylation method in mapping unpaired cytosines at the sequence level are distinct advantages. The method complements others using single-strand-specific nucleases or a combination of chemical and nuclease treatments (7, 23), in terms of both selecting reaction conditions and avoiding artifacts that might be introduced by the use of structural probes. In addition, the cleavage of the DNA chain at the N^3-methylcytosine positions is relatively insensitive to the nucleotide sequence, compared to the sequence dependence of cleavages by enzymes.

Our studies on the RNA polymerase–promoter complexes demonstrate the applicability of the cytosine methylation reaction to probing the positions of unpaired cytosines in protein–DNA complexes and add further evidence supporting the RNA polymerase-induced unpairing of bases in a specific region of the promoters (8, 9).

The sharp temperature dependence of the unpairing of the cytosines in the *lac* UV5 and *tac* promoter complexes of RNA polymerase suggests that the strong temperature dependence of the formation of an initiation complex, in which RNA chain initiation occurs, is closely related to the RNA polymerase-induced unpairing of bases in the promoter. The three promoters we examined show rather distinct differences in their uncoiling by the polymerase. Whereas unpairing of bases in the *lac* UV5 and *tac* promoters is readily detected by cytosine methylation, no cytosine in either strand of the *trp* is made reactive to dimethyl sulfate by polymerase binding.

The −35 region of the *trp* promoter is identical in sequence to the −35 region of the *tac* promoter and corresponds to the canonical sequence for this region deduced from a compilation

Transcription Initiation