Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 1301–1305, February 1991
Biochemistry

Phased cis-acting promoter elements interact at short distances to direct avian skeletal α-actin gene transcription

(trans-acting factors/gene regulation/torsional phasing/DNA twisting)

KING-LAU CHOW*†, MICHAEL E. HOGAN‡, AND ROBERT J. SCHWARTZ*§

*Department of Cell Biology, Baylor College of Medicine, Texas Medical Center, Houston, TX 77030; and ‡Center for Biotechnology, Baylor College of Medicine, 4000 Research Forest Drive, The Woodlands, TX 77381

Communicated by Donald M. Crothers, October 10, 1990 (received for review August 3, 1990)

ABSTRACT **Recently, site-directed mutagenesis uncovered four positive cis-acting elements in the 5′ promoter region of the chicken skeletal α-actin gene that directs myogenic tissue-restricted expression. In this study, interactions between the four promoter sites were examined by means of a series of insertion mutations that increased the linker region between adjacent elements by roughly half or complete DNA helical turns. Unexpectedly, transcriptional activity for all three sets of linker mutants, as assayed with a chloramphenicol acetyltransferase reporter gene, was found to vary in a fashion resembling a damped sinusoid with a period of roughly 10 base pairs, where the sinusoidal maxima appeared when length was increased by half-integral number of helix turns. We present a model which states that in the undistorted wild-type 5′ flanking sequence, linker domains position each of the four promoter sites on the helix face opposite that of its immediate neighbors; when any of the three linkers is increased by approximately a half-integral number of helix turns, pairs of neighboring promoter sites are brought into alignment. We propose that this is the required orientation for inducing skeletal muscle-specific promoter activity, achieved in the wild-type promoter as a result of protein-induced torsional deformation.**

A eukaryotic transcription unit usually consists of a promoter with a combination of multiple regulatory elements (1, 2). Interactions between the various regulatory elements are thought to be mediated by protein–protein contacts. In the case of enhancers or silencers, such interactions appear to be exercised in a fashion that is not dependent upon the exact distance between interacting sites (3). That behavior is to be compared with that of distal promoter elements and some upstream activating sequences, which display a critical dependence upon separation (4, 5) and appropriate alignment for efficient transcription.

At least three classes of pairwise interactions have been identified to explain such a detailed length and spatial dependence. (*i*) Several examples, including the λ repressor (6), *araCBAD* regulatory region (7), and simian virus 40 early promoter (8), favor the mutual interaction of paired elements in a configuration that allows proteins to bind to the same face of the DNA helix. In this way, promoter elements separated by relatively short distances interact by means of protein–protein contacts that bend the helix and potentiate transcription. (*ii*) Less favored is the interaction between paired elements positioned on opposite sides of the DNA helix. Only one example, involving the frog ribosomal gene promoter, was recently cited (9). In this type of stereospecific orientation, restoration of alignment between promoter elements stimulated transcription and altered the sites of initiation. Thus, the torsional order of a helix might restrict promoter activity mediated through protein–protein contacts with respect to deformation, or twisting, of the DNA helix. (*iii*) In several instances, the alignment of DNA binding sites was shown not to play a strict role in promoter activity (10, 11) and may relate in part to the lateral flexibility of the binding proteins.

Our recent study (12) revealed that a combination of four positive cis-acting elements in the chicken skeletal α-actin gene promotor is required for efficient transcription in cells. These regulatory motifs are described as the upstream T+A-rich element [UTA, base pairs (bp) −170 to −190], upstream CCAAT box-associated repeat (U-CBAR, bp −123 to −138), downstream CBAR (D-CBAR, bp −94 to −80), and the ATAAAA box (bp −25 to −30). They are positioned in the promoter in a symmetric manner with respect to a center of twofold symmetry at bp −108 (13). In the work described here, interactions between these actin promoter elements was evaluated by increasing the size of the linker DNA segment between each pair of elements. This analysis has revealed that actin promoter activity is sensitive to the size of all three linker domains. One important corollary of this analysis is that, analogous to DNA bending, the undeformed wild-type actin promoter is not optimized with respect to interaction between adjacent sites but may require approximately one-half turn of torsional deformation for each of the three pairwise interactions throughout the α-actin promoter.

MATERIALS AND METHODS

Engineering of the Promoters. The MCAT clone (12) contains the 5′ flanking region of the chicken skeletal α-actin gene from bp −421 to bp +24 linked to the bacterial chloramphenicol acetyltransferase (CAT) gene. Insertional mutations were positioned at synthetic *Bgl* II sites that were previously shown to function as neutral regions in the promoter, such as between D-CBAR and ATAAAA in M8 (series I), between U-CBAR and D-CBAR in M19 (series II), and between UTA and U-CBAR in M26 (series III). Mutant promoters containing 4-nucleotide insertions were prepared in each series by *Bgl* II digestion, filling in of "sticky" ends by Klenow DNA polymerase, and blunt-end-ligation. Complete-DNA-helical-turn insertion mutants were constructed in *Bgl* II-digested and dephosphorylated promoters. Ten-base-pair *Bgl* II linkers were ligated in tandem, recut with *Bgl* II to yield 10-mer linkers with protruding ends, and then ligated at a high molar ratio to the pretreated promoter. Additional, half-DNA-helical-turn insertions were generated by ligation of 10-bp *Bgl* II linkers with the prepared blunt-ended promoters that contain 4 extra nucleotides.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: CAT, chloramphenicol acetyltransferase; CBAR, CCAAAT box-associated repeat; U-CBAR, upstream CBAR; D-CBAR, downstream CBAR; UTA, upstream T+A-rich element.

†Present address: Department of Molecular Genetics, Albert Einstein College of Medicine, Yeshiva University, 1300 Morris Park Avenue, Bronx, NY 10461.
§To whom reprint requests should be addressed.

Transfection into Cell Cultures. Mutant promoter–CAT constructions (6 μg of plasmid) were cotransfected with pRSV-β-Gal (2 μg), an internal standard for normalizing transfection assays, into chicken embryonic myoblasts 24 hr after plating as described (12). The cells were harvested 72 hr later and assayed for CAT and β-galactosidase activity (13). Comparison of the standardized CAT activity with the wild-type construction provided a percentage of wild-type activity.

DNase I Footprint Analysis. A nuclear extract from 17-day embryonic chicken muscle (14) was used for DNase I footprint analysis. The skeletal α-actin promoter in MCAT was cut with *Hin*dIII, asymmetrically labeled at the 3′ end of the coding strand (position +24) by Klenow fill-in with [α-^{32}P]dATP (3000 Ci/mmol; ICN; 1 Ci = 37 GBq), and released from the vector by *Sma* I digestion (position −201). The probe (10^4 cpm/ng) was mixed with 30 ng of poly(dI-dC) and was preincubated at room temperature in 10% (vol/vol) glycerol/10 mM Hepes, pH 7.9/20 mM KCl/4 mM $MgCl_2$/0.1 mM EDTA/0.25 mM dithiothreitol with nuclear extract (4 μg of protein). DNase I (2 units) was added to the reaction mixture (final volume, 25 μl) for 2 min. Reactions were stopped by the addition of phenol and chloroform. Deproteinized DNA was heat-denatured in 95% formamide with 20 mM EDTA and loaded onto a 7 M urea/6% polyacrylamide gel. After electrophoresis the gel was dried and autoradiographed with Kodak XAR-5 film.

RESULTS AND DISCUSSION

Increased Distance Between CBAR Elements Gives Rise to an Oscillatory Relationship Between Length and Activity. Insertion mutations generated at the center of the linker domain between CBAR elements were analyzed by the transient-expression assay system. Included in the analysis were five insertions that increased length by 10*n* or 10*n* + 4 bp between the paired CBARs. In Fig. 1, the activity of the mutants has been catalogued as a function of increasing linker length. The data are well described as a damped sinusoidal decay with maxima at increments of 4, 14, and 24 bp. Minima are detected at 10, 20, and 30 bp. Beyond 30 bp the phase of the pattern becomes altered. However, because transcriptional activity has decreased ≈100-fold at that point, emphasis is placed upon the six data points that display activity significantly greater than that of the promoterless control (see Fig. 1*A*).

That promoter activity varies as much as 50-fold as a function of linker insertion requires that either protein–DNA or protein–protein contacts have been altered by the length increase. In planning these mutations, a site was chosen that did not inhibit promoter function or interfere with factor binding to adjacent regulatory elements. In this regard, there is no evidence for protein–DNA contacts within this inter-CBAR domain that are required for promoter activity. By this criterion, we favor the idea that insertional mutagenesis has altered the spatial relation between adjacent CBAR factor binding sites. The observed oscillatory relation between activity and length (with roughly a 10-bp period) suggests a physical origin that is dependent upon the pitch of the helix. The surface orientation of helix sites displays such a dependence.

Therefore an adequate rationalization of the observed oscillation for the CBAR–CBAR linker is as follows. The two adjacent CBAR sites are separated by 45 bp (4.5 helix turns), center to center, in the wild-type promoter. If the helix were to remain undistorted, identical protein binding interactions at CBAR sites would position neighboring protein–DNA complexes on opposite faces of the helix, which is presumed to be undesirable. It is proposed that a 4-bp linker increase brings CBAR binding sites (now 49 bp center to center) onto the same helical face, thereby allowing for protein–protein contacts. A length increase of 14 or 24 bp induces an equivalent favorable torsional orientation. However, as the linker increases, the two binding sites become separated, thereby requiring bending of the duplex to achieve protein–protein contact. As a result of that additional bending requirement, the stability of the complex is reduced, giving rise to the observed damping of transcriptional activity with length increase.

DNase I-Hypersensitive Sites Are Dependent on the Orientation of CBAR Elements. The above model predicts that upon binding CBAR factors, the wild-type promoter may become torsionally stressed, whereas a mutant with a 4-bp insertion will not experience such deformation. DNase I footprinting was employed to explore possible alteration of DNA structure in the CBAR–CBAR linker after binding muscle nuclear proteins. The wild-type promoter showed a DNase I-hypersensitive site (marked with an asterisk) in the linker domain between the two CBARs (Fig. 2). When a promoter with a 4-bp linker increase (M19+4) was footprinted, the hypersensitive site was clearly eliminated, whereas two new, weaker hypersensitive sites were identified in the M19+10 mutant (10-bp linker increase). DNase I footprinting within the U- and D-CBAR indicated that direct interaction of nuclear binding proteins with the CBAR sites was not altered by insertional mutagenesis within the linkers (Fig. 2). In addition, analysis of the TAB1 mutant (Fig. 2), in

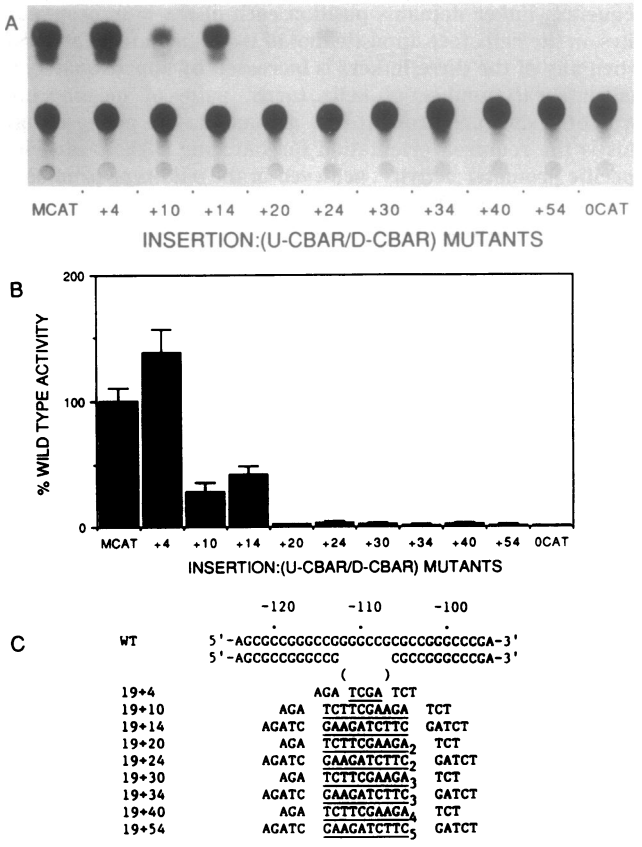

FIG. 1. An increase of the linker between the D-CBAR and U-CBAR gives rise to transcriptional activity that varies with a 10-bp period. Insertion mutations were generated in M19CAT at bp −109 (*C*), as described in *Materials and Methods* and transfected into primary myoblast cultures. CAT activity in cell lysates was assayed by conversion of radiolabeled chloramphenicol to its acetylated derivatives (autoradiograph of thin-layer chromatogram in *A*; quantification in *B*). Error bars in *B* represent standard deviations about a mean of at least four independent assays. 0CAT, promoterless control; WT, wild type.

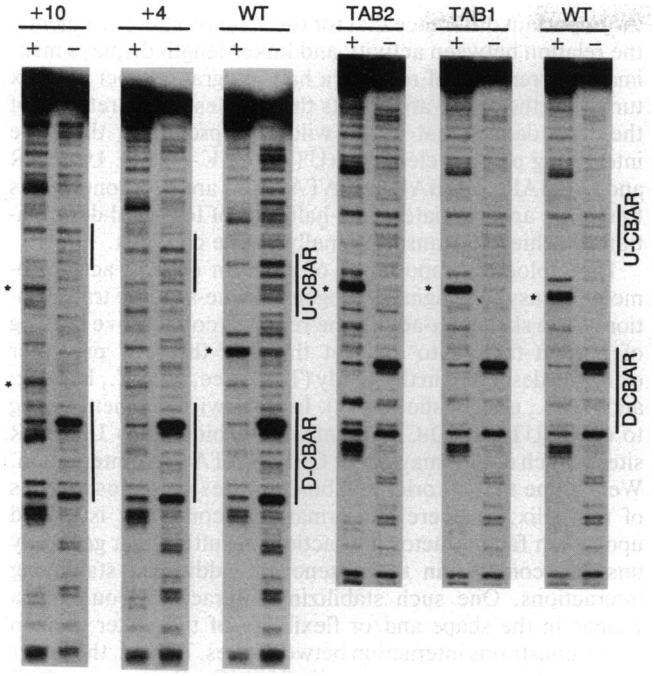

FIG. 2. DNase I-hypersensitive sites are altered by an increase of roughly half or complete helical turns of the CBAR–CBAR linker. Lanes +, crude extract; lanes −, control. (*Left*) In the M19+4 mutant, the enhancement band was reduced, while two new hypersensitive sites appeared in the M19+10 promoter mutant. (*Right*) Substitution of AT for GC in the TAB1 and TAB2 mutants. Hypersensitive sites are marked by asterisks. WT, wild type.

which an A was substituted for G at position −110, and the TAB2 mutant, in which AT was substituted for GC at −110 and −109, suggests that disruption of the G+C-rich linker by A·T base pairs does not, by itself, eliminate DNase hypersensitivity. The above observations are consistent with the idea that binding of nuclear protein factors to the CBAR sites induces a conformational change that facilitates DNase I binding and cleavage within the wild-type CBAR-CBAR linker but not within the linker with a 4-bp insertion.

Increased Distances Between D-CBAR and ATAAA and Between U-CBAR and UTA Elements Give Rise to an Oscillatory Pattern of Activity. Previous experiments suggested that the ATAAAA box must interact with the other upstream promoter elements to potentiate transcription (12). Seven insertional mutations were generated at the *Bgl* II site in M8CAT, at bp −43 of the promoter, 20 bp from the end of the ATAAAA element, and assayed for transient expression. Insertional mutagenesis of the CBAR–ATAAAA linker gave rise to a damped sinusoidal activity dependence with well-defined maxima at length increases of 4, 14, 24, and 34 bp and minima at 10, 20, and 30 bp (Fig. 3). The simplest interpretation of the oscillatory activity–length dependence is that protein factors which bind to the ATAAAA box also interact directly with factors bound to the D-CBAR domain. As for the hypothetical CBAR–CBAR interation, the fact that maxima are observed when helix length is increased by roughly one-half turn, and minima at multiples of one turn, suggests that in the undistorted wild-type promoter, the two binding sites are positioned upon opposite faces of the duplex and are brought into torsional register due to protein–protein contacts.

Linker insertion mutations between the UTA element and the U-CBAR element were prepared and analyzed by the CAT assay (Fig. 4). Again, an oscillatory pattern of diminishing promoter activity was obtained, with local maxima at 4, 14, and 24 bp and minima at 10 and 20 bp of linker increase.

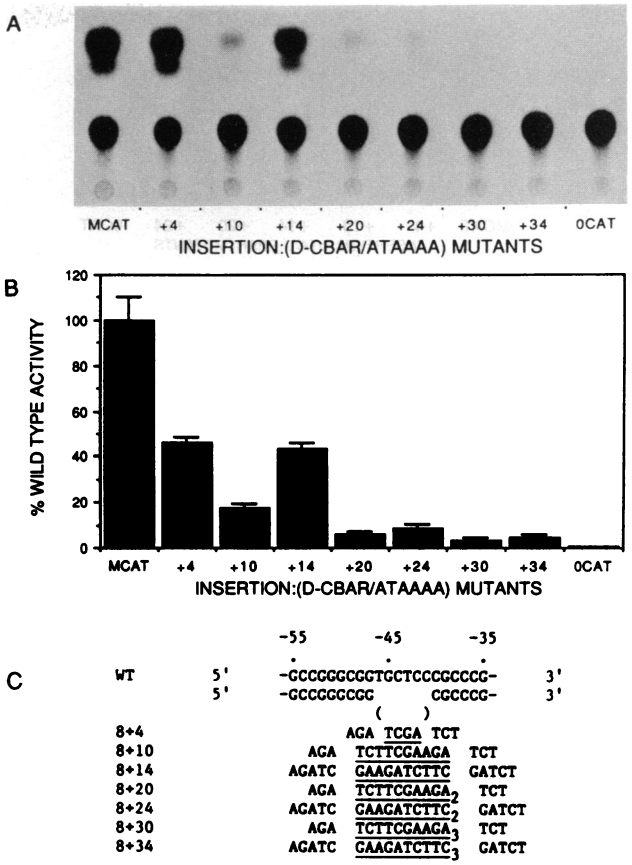

FIG. 3. An increase of the linker between D-CBAR and ATAAAA gives rise to cyclical transcriptional activity that varies with a 10-bp period. Insertion mutations were constructed at bp −43 of M8CAT. See Fig. 1 legend for further explanation.

As for the other two adjacent pairs, these data suggest that the UTA and U-CBAR elements can interact (presumably due to protein contacts) and that in the wild-type promoter, the two sites are oriented upon the opposite sides of the helix, requiring torsional deformation (or one-half turn of linker increase) to bring the sites into register.

CONCLUSIONS

Hypothesis: Protein Binding Sites Within the α-Actin Promoter Are Positioned so as to Require Torsional Deformation of the Helix. In the λ phage repressor system, cooperative binding of the repressor molecule to the operator sequences occurs only when the sequences are separated by complete DNA helical turns, not by half turns (6). Similar stereospecific alignment has also been demonstrated in the binding of AraC protein at the *araI* and *araO2* sites of the *araCBAD* regulatory region (7, 15). Repression of the *araBAD* promoter transcription was normal for insertion of a complete DNA helical turn between the two elements but was inhibited by insertion of a half helical turn. In eukaryotes, phasing effects have also been reported for interaction between domain A of the enhancer and the 21-bp repeat and between the 21-bp repeat and the normal TATA box in transcriptional regulation of the simian virus 40 early promoter (8). The simplest interpretation of those data is that in the wild-type promoter, adjacent protein factors are positioned on the same side of the helix, so that a half-integral linker increase serves to position the pair in an unfavorable relative position.

As shown in Figs. 1–4, sequence manipulations between regulatory sites resulted in a pattern of transcriptional activity that is best described as a damped sinusoid with a period

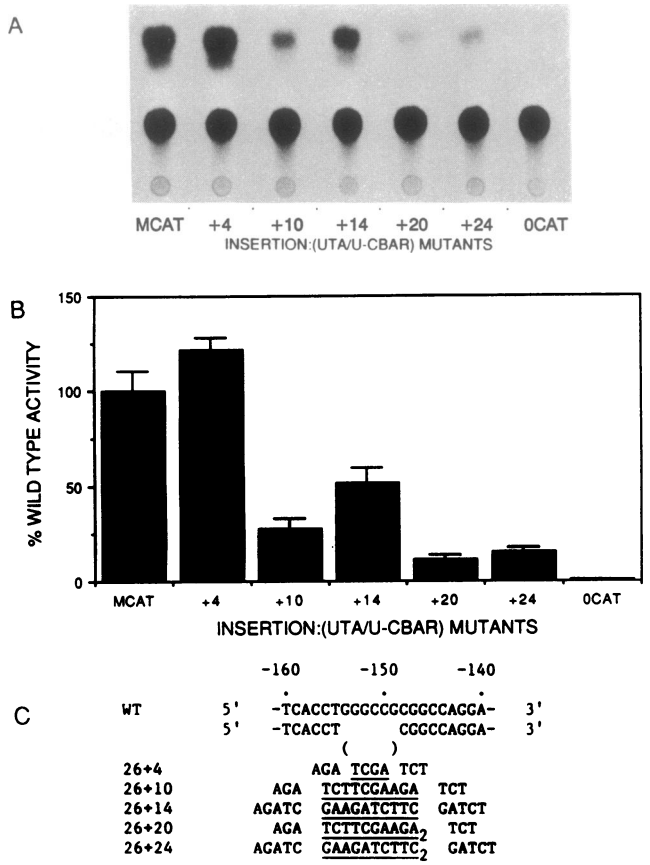

FIG. 4. An increase of the linker between the UTA and U-CBAR gives rise to transcriptional activity that oscillates with a 10-bp period. Insertion mutations were generated at bp −151 of M26CAT. See Fig. 1 legend for further explanation.

of ≈10 bases. The form of the oscillatory activity–structure relation is similar to the above-mentioned examples, but with the important difference that for the actin promoter elements, the relation between activity and linker length displays maxima at increments of roughly a half-integral number of helix turns. By the above arguments the simplest interpretation of the actin data is that in the wild-type promoter, the three interacting pairs of elements (UTA and U-CBAR, U-CBAR and D-CBAR, D-CBAR and ATAAAA) are positioned so as to require approximately one-half turn of torsional deformation to achieve a transcriptionally active complex.

The biological importance of such an orderly actin promoter is easy to imagine. Myogenic tissue-specific transcription of the skeletal α-actin gene requires cooperative binding of protein factors to at least three of the four promoter elements described in this study (T. C. Lee, K.L.C., P. Fang, and R.J.S., unpublished work). In a pairwise manner binding to either UTA or D-CBAR influences binding to U-CBAR sites, which in turn may affect CBAR–ATAAAA interaction. We assume that by orienting binding sites on opposite sides of the helix, a severe conformational constraint is placed upon such factor–factor interaction, resulting in a generally unstable complex in the absence of additional stabilizing interactions. One such stabilizing interaction would be a change in the shape and/or flexibility of the linker domain which constrains interaction between sites. Indeed, the linker regions of the actin promoter display a G+C content > 90% (12), which suggests that such DNA may be unusually stiff with respect to bending and/or torsional deformation (16). Substitution of A+T-rich sites within these G+C-rich sequences stimulated developmentally timed transcriptional activity of the α-actin promoter (12), perhaps due to an increase of linker flexibility in the more A+T-rich mutants. Rather than by sequence manipulation, it is interesting to consider that α-actin expression could be manipulated by the binding of protein factors or other molecules to the α-actin promoter linker domains. Like a sequence-dependent flexibility change, such auxiliary factors might work by altering linker twist and curvature by means of direct binding to linker DNA or by stabilizing the contacts formed by UTA, CBAR, and ATAAAA factors.

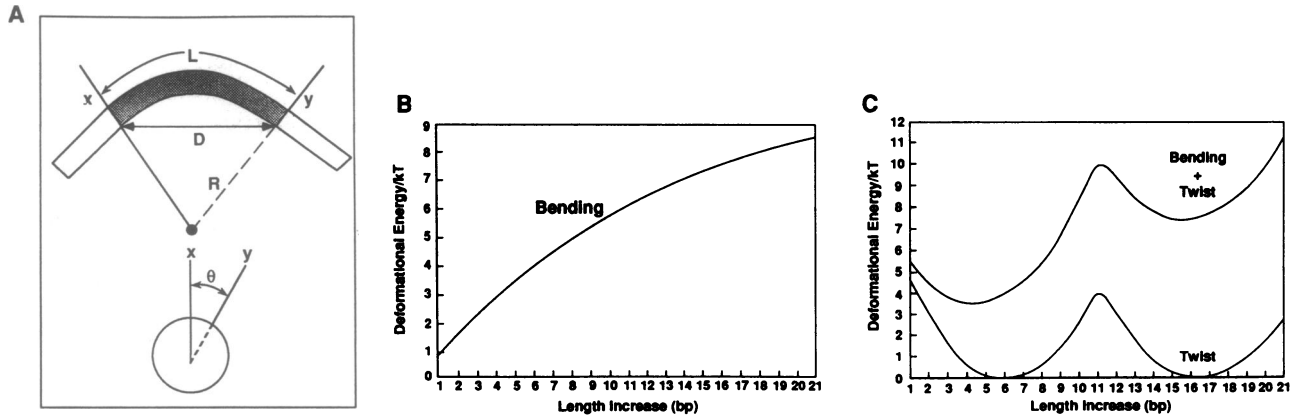

FIG. 5. The stress imposed upon the DNA linker by interacting protein-factor binding sites: A simple elastic model. The DNA helix can be modeled as a stiff elastic coil (17). If the ends of a DNA linker domain with length L are constrained to a separation D and an angular orientation θ due to factor binding, the radius of curvature of the resulting smooth DNA curve is adequately described by $R = \sqrt{L^3/24(L - D)}$. Such a dependence is displayed in *A*. Simple elastic theory predicts that the bending work required to achieve such curvature is given by $U_{bend} = EIL/2R^2$, where I is a constant (2.3×10^{-28} cm^4) and E is the Young's modulus for the DNA sequence under consideration (≈1×10^9 dyne·cm^{-2} for random-sequence DNA; ref. 16). In *B*, we have plotted the bending work for a constrained linker domain relative to thermal energy kT, as a function of linker contour length L and assuming, for example, D = 46 bp × 3.35 Å. If analogous to D there is an angle θ that describes the required phase orientation between sites x and y in the bond complex, then $\theta = (2\pi \text{ rem } L/P - \phi)$, where rem L/P is the phase angle between sites x and y in the undistorted linker and ϕ is the angular distortion required to bring sites x and y into register. The work required to twist an elastic rod is $U_{twist} = IG\phi^2/2L$, where G is the shear modulus of the helix and is approximately $E/3$ for random-sequence DNA (16). In *C*, lower curve, we have combined the above relations: $U_{twist} = (I/2L)(2\pi \text{ rem } L/P - \theta^2)$ and have presented the calculated torsional energy for an unbent DNA linker segment, assuming that sites x and y are parallel in the bound complex ($\theta = 0$) and a helix pitch of 10.4 bp. In *C*, upper curve, we have combined the relations for bending and twisting in *B* and *C* to yield a curve that describes the bending and torsional energy required to accommodate a specified linker length change. For these calculations we have assumed D = 46 bp and $\theta = 0$.

The Observed Transcriptional Activity Can Be Modeled with a Simple Elastic Coil Model. It can be shown that the sinusoidal form of the observed length–activity dependence can be adequately interpreted in terms of geometry and the known elastic properties of a DNA helix. In Fig. 5, such an analysis is presented in terms of the simplest model, one which considers only the interaction between nearest neighbor sites.

If transcription requires that two regulatory elements on a helix are held at a fixed distance D and at a fixed angular orientation (by contacts made between their respective protein factors), then the work required to attain the transcriptionally active state becomes a function of the resistance offered by the intervening DNA linker domain, which is a function of DNA shape and DNA flexibility parameters (Fig. 5*A*). If the length L of the linker region is increased, the ratio D/L will decrease, which specifies that an arc is created with a radius of curvature which decreases as L is increased. Since the DNA helix is adequately described as a stiff elastic body (17), then as the radius of curvature decreases, the work required to form the transcriptionally active complex will increase. If transcription is a function of that apparent binding energy, activity will decrease continuously with increasing linker separation L (Fig. 5*B*), as is seen for each of the three sets of linker mutants.

Independent of bending considerations, if protein factors interact to form a well-defined complex, that complex can also be described by an orientation angle θ, such that if θ equals zero, the interacting sites are positioned on the same face of the helix in the bound complex, whereas a θ value of π corresponds to a bound complex with protein monomers on opposite faces of the helix. As L is increased, the angular orientation of adjacent sites in the undistorted helix will move into, and then out of, register with a period of $\approx$10 bases. When out of register, the helix must be deformed by twisting to form the active complex. When in register, torsional distortion is not required and less work is required to form the complex. The combined effects of length and torsional phase can be predicted by elastic theory and are presented in Fig. 5*C* (lower curve). When both torsional and bending energetics are considered (Fig. 5*C*, upper curve), the measured data (Figs. 1–4) can be described by a model in which an increase of the linker between protein binding domains gives rise both to bending (damping of activity with length) and to torsional distortion of the promoter domain (which produces the oscillatory character of the data). A quantitative analysis of that type, based upon independently measured DNA flexibility parameters (to be presented elsewhere), suggests that the activity of each of the three sets of linker mutants is adequately described by such a simple model.

We acknowledge the helpful advice of R. E. Austin. This work was supported by U.S. Public Health Service Grants HL38401 and 87-HL-26-H to R.J.S. and a grant from the National Cancer Institute to M.E.H.

1. McKnight, L. L. & Tjian, R. (1986) *Cell* **46,** 795–805.
2. Maniatis, T., Goodbourn, S. & Fischer, J. A. (1987) *Science* **236,** 1237–1245.
3. Gralla, J. D. (1989) *Cell* **57,** 193–195.
4. Zenke, T., Grundstrom, T., Matthes, H., Wintzerith, M., Scharz, C., Wilderman, A. & Chambon, P. (1986) *EMBO J.* **5,** 387–397.
5. Goodbourn, S., Burstein, H. & Maniatis, T. (1986) *Cell* **45,** 601–610.
6. Hochschild, A. & Ptashne, M. (1986) *Cell* **44,** 681–687.
7. Lee, D.-H. & Schleif, R. F. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 476–480.
8. Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. & Chambon, P. (1986) *Nature (London)* **319,** 121–126.
9. Pape, L. K., Windle, J. J. & Sollner-Webb, B. (1990) *Genes Dev.* **4,** 52–64.
10. Ruden, D. M., Ma, J. & Ptashne, M. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4262–4266.
11. Wirth, T., Stradt, L. & Baltimore, D. (1987) *Nature (London)* **329,** 174–178.
12. Chow, K.-L. & Schwartz, R. J. (1990) *Mol. Cell. Biol.* **10,** 528–538.
13. Grichnik, J. M., French, B. A. & Schwartz, R. J. (1988) *Mol. Cell. Biol.* **8,** 4587–4597.
14. Dignam, J. D., Martin, P., Shastry, B. S. & Roeder, R. G. (1983) *Methods Enzymol.* **101,** 582–598.
15. Dunn, T. M., Hahn, S., Ogden, S. & Schleif, R. F. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 5017–5020.
16. Hogan, M. E. & Austin, R. E. (1987) *Nature (London)* **329,** 263–266.
17. Barkley, M. D. & Zimm, B. H. (1979) *J. Chem. Phys.* **70,** 2991–3007.