The EMBO Journal vol.6 no.5 pp.1513–1522, 1987

Long-range interactions of multiple DNA structural transitions within a common topological domain

Michael J.Ellison, Matthew J.Fenton, Pui Shing Ho and Alexander Rich

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Communicated by A.Rich

Local DNA conformations that are underwound with respect to the right-handed B form are favored in negatively supercoiled DNA. However, when multiple transitional events coexist within a common topological domain, they must compete with one another for the available free energy of negative supercoiling. Recently we developed a general theoretical model capable of predicting the behavior at equilibrium of defined sequences in a variety of competitive situations. In the present work we have applied this theory to predict the formation of Z-DNA as a function of superhelicity in stretches of $d(CG)_m$ and $d(CA)_n$ when they are forced to compete with one another in the same plasmid. The observed behavior of these competing sequences is in close accord with theoretical predictions. These results indicate that sequences separated by large distances can effect the transitional behavior of each other in a complex manner which is independent of the relative orientation of the participating segments. The pattern of transitional events is strongly dependent on levels of DNA supercoiling, ambient conditions and on the nature and number of the sequences involved. Although in the present work we apply the model specifically to the Z-DNA conformational transition, the results of this study may have general relevance to a variety of biological processes in which the helical repeat of DNA is reversibly altered, including the initial steps in transcription, replication and recombination.

Key words: Z-DNA/B-DNA/supercoiled DNA/long-range interactions/B–Z transition

Introduction

A considerable number of *in vitro* and *in vivo* studies have strongly suggested that modulations in the level of DNA supercoiling can give rise to profound and often contradictory biological effects. In prokaryotes the mechanism for regulating levels of DNA superhelicity is established largely by the antagonistic actions of DNA gyrase which negatively supercoils DNA, and topoisomerase 1 which relaxes negatively supercoiled DNA. In eukaryotes the mechanism by which superhelicity is regulated is not yet clearly defined (for reviews see Gellert, 1981; Wang, 1985). Recent studies have shown that the initiation of DNA replication *in vitro* from an *Escherichia coli* origin requires a negatively supercoiled template (Baker *et al.*, 1986). In addition, site-specific integration of phage λ DNA into the *E. coli* chromosome appears to occur most efficiently when the λ DNA molecule is negatively supercoiled (reviewed by Weisberg and Landy, 1983). On the other hand the effect that superhelicity has on transcription is considerably more complicated. For some genes, transcription is enhanced by negative supercoiling while other genes, such as DNA gyrase, become suppressed (reviewed by Wang, 1985). These types of observations have strengthened the viewpoint that the regulation of DNA supercoiling may be involved in controlling a number of complex cellular events.

While the structural details that underly each of these supercoiling-dependent processes are poorly understood, they can none the less be loosely rationalized in terms of the dynamic properties peculiar to negatively supercoiled DNA. Negative supercoiling is associated with a positive free energy change that can be utilized to drive any process that involves helix unwinding. Transcription, replication and recombination are processes that involve strand separation and as such may derive part of their energetic requirements from DNA templates that are negatively supercoiled.

Superimposed on this general effect, however, is the potential for specific and differential genetic effects which could arise when regulatory sequences undergo supercoiling-induced conformational changes (Wang, 1982). A number of underwound structures have been shown to form spontaneously in negatively supercoiled DNA in which base-pairing interactions are largely preserved. The nature of these transitions and the ease with which they form are heavily dependent on sequence and the magnitude of helical unwinding associated with each. While the potential for a great many structural alternatives probably exists for negatively supercoiled DNA, cruciforms (reviewed in Sullivan and Lilley, 1986) and left-handed Z-DNA (reviewed by Rich *et al.*, 1984) remain the most extensively characterized examples to date. Both of these structures have been demonstrated to exist *in vivo* under special circumstances (Haniford and Pulleyblank, 1983a, 1985). The possibility that supercoiling-induced transitions in DNA can alter the binding characteristics of proteins that activate or repress specific genetic functions naturally expands the complexity of events brought into play by changes in superhelicity.

Matters become further complicated when it is recognized that discrete, locally underwound transitions occurring within the same topological domain must compete with one another for the limited free energy of negative supercoiling. In linear or nicked DNA molecules, conformational transitions arising in different parts of the molecule occur independently of one another due to the presence of one or more swivel points. Closed circular DNA molecules, however, are topologically constrained and therefore different conformations must share the available free energy of supercoiling between them. Each conformational event, then, becomes contingent on other transitional events elsewhere in the molecule. The interplay between transitional events in closed circular DNA was first recognized by Benham (1981, 1982) as a potentially complex process. While these studies provided cosiderable insight into the behavior competing transitions, the formalism developed by Benham could not be readily tested experimentally.

Recently we have developed a generalized quantitative model that has allowed us to predict how multiple discrete transitions occurring within defined sequences will compete with one another

within the same superhelical domain (Kelleher *et al.*, 1986). In the present work the model is tested experimentally by examining the competition for Z-DNA formation in the sequences $d(CG)_m$ and $d(CA)_n$ when both sequences are situated noncontiguously within the same plasmid molecule. The theoretical and experimental results presented here illustrate that complex long-range transitional patterns can be established in negatively supercoiled DNA that change as the level of supercoiling is varied. While the example presented here relates to competing segments that form Z-DNA, the general problem is relevant to any set of competing processes that unwind the DNA. How these competitive patterns may contribute to the temporal programming of genetic events is discussed.

Results

The model

To date, the polymers $d(CG)_m$ and $d(CA)_n$ remain the best structurally and thermodynamically characterized examples of sequences known to undergo a supercoiling-induced transition. In the present work, therefore, we have chosen to examine the behavior of these sequences when forced to compete against one another for the free energy of negative supercoiling.

The extent of Z-DNA formation in plasmids containing single stretches of $d(CG)_m$ or $d(CA)_n$ as a function of negative superhelicity can be approximated by a simple statistical mechanical formulation of the zipper model (Peck and Wang, 1983; Vologodskii and Frank-Kamenetskii, 1984). In this model, Z-DNA is presumed to nucleate at some position within the sequence by the formation of the B−Z junctions which then propagate Z-DNA by increments until the boundaries of the sequence are reached. Each of these two events has associated with it an energetic penalty and each result in a defined degree of unwinding of the B double helix. Transition in these sequences is facilitated by the positive free energy of negative supercoiling which varies as the square of the plasmid linking difference from the relaxed position to moderate levels of superhelicity (Depew and Wang, 1975; Pulleyblank *et al.*, 1975). A more complete description of this model and its underlying assumptions as it relates to the present work has been discussed previously (Ellison *et al.*, 1985, 1986; Kelleher *et al.*, 1986).

Five parameters have been shown to be sufficient for describing with reasonable accuracy the unwinding characteristics of $d(CG)_m$ and $d(CA)_n$ as a function of the changing linking number: (i) the free energy of negative supercoiling associated with each topoisomer, ΔG_τ; (ii) the unwinding resulting from the conversion of a dinucleotide from B-DNA to Z-DNA; (iii) the unwinding occurring at the B−Z junctions; (iv) the energy required to stabilize a dinucleotide in the Z form, ΔG_{B-Z}; and (v) the energy necessary for the stabilization of the B−Z junctions, ΔG_j. These values have been experimentally determined for $d(CG)_m$ (Peck and Wang, 1983& and $d(CA)_n$ (Vologodskii and Frank-Kamenetskii, 1984) and are listed in the legend to Figure 1.

Two considerations allow for the extension of the above model to situations where multiple and discrete transitions occur within the same topological domain. For truly independent transitions, the probability of events occurring simultaneously in different sequences is simply the product of their individual probabilities. In addition, it must be realized that the free energy of supercoiling available for transitional processes throughout the plasmid must be partitioned between each of the transitions permitted to occur. A formal description of this model has been provided elsewhere (Kelleher *et al.*, 1986). It is noted that the multiple transition model involves no additional parameters or assumptions other than those that already pertain to the single transition case.

We have illustrated this competition theoretically by placing the two sequences $d(CG)_8$ and $d(CA)_n$ [where the length of the $d(CA)_n$ segment is permitted to vary from $n = 0$ to $n = 60$ dinucleotides] within the context of a 4.3-kb plasmid. Both Z-forming segments are imagined to be separated from one another by a considerable stretch of B-DNA which is long enough to prevent their direct structural interaction. Figure 1A shows the predicted unwinding resulting from Z-DNA formation in the sequence $d(CG)_8$ both as a function of the negative superhelicity of the plasmid and as a function of the length of its competitor, $d(CA)_n$. The unwinding arising from Z-DNA formation is calculated from the observation that the conversion of 1 bp from the B form to the Z form alters the helical pitch by −0.18 turns (Wang *et al.*, 1979) and from the estimated −0.8 turns contributed by the B−Z junctions (Peck and Wang, 1983). Thus the complete conversion of $d(CG)_8$ to the Z form is expected to produce 3.7 turns total unwinding (16 bp × 0.18 turns/bp + 2 × 0.4 turns/junction). When the length of the competing $d(CA)_n$ stretch is set at zero, the unwinding in $d(CG)_8$ increases sigmoidally to a plateau of ~3.7 turns, in accord with previous experimental findings (Peck and Wang, 1983). As the length of the $d(CA)_n$ increases, the behavior of the $d(CG)_8$ segment undergoes a complex modulation with increasing values of the negative linking difference. At intermediate values of the linking number, a trough becomes discernible which culminates at longer lengths of $d(CA)_n$ in a shift of the B−Z equilibrium from a preponderance of Z-DNA to B-DNA (Figure 1A). In comparison to $d(CG)_8$, the behavior of $d(CA)_n$ in this coupled system is less complex (Figure 1B). As the length of the $d(CA)_n$ stretch is increased from 0 to 60 dinucleotides, the transition profile of this sequence changes from a sigmoidal curve to one with subtle inflections. In addition, the maximum degree of unwinding observed for $d(CA)_n$ increases with increasing sequence length.

Testing the model

The most striking feature of the predicted competition between $d(CG)_8$ and $d(CA)_n$ is the complex Z-forming behavior of the shorter $d(CG)_8$ segment when forced to compete against significantly longer stretches of $d(CA)_n$. In order to examine the extent to which this effect actually occurs, two plasmids were constructed, each conforming to the general constraints set down for the model presented in Figure 1. Plasmid pZ2a, for example, is 4.3 kb in size and contains a $d(CG)_8$ segment separated ~600 bp from a 60-bp stretch of alternating $d(CA)_n$ (Figure 2A). Plasmid pZ2b is identical to pZ2a with the exception that the length of $d(CA)_n$ is ~120 bp in length (Figure 2B). The distance separating both Z-forming segments is presumed to be great enough to allow for the independent behavior of each sequence.

The observation that negative supercoiling stabilizes Z-DNA (Singleton *et al.*, 1982) and that the extent of the transition can be monitored by agarose gel electrophoresis (Klysik *et al.*, 1981) has provided a useful tool for studying the dynamics of supercoiling-induced transitions in DNA. The most effective means of analyzing such transitions is by the two-dimensional electrophoresis of plasmid topoisomers (Haniford and Pulleyblank, 1983b; Peck and Wang, 1983). Topoisomers which vary in their linking difference also vary from one another in their degree of

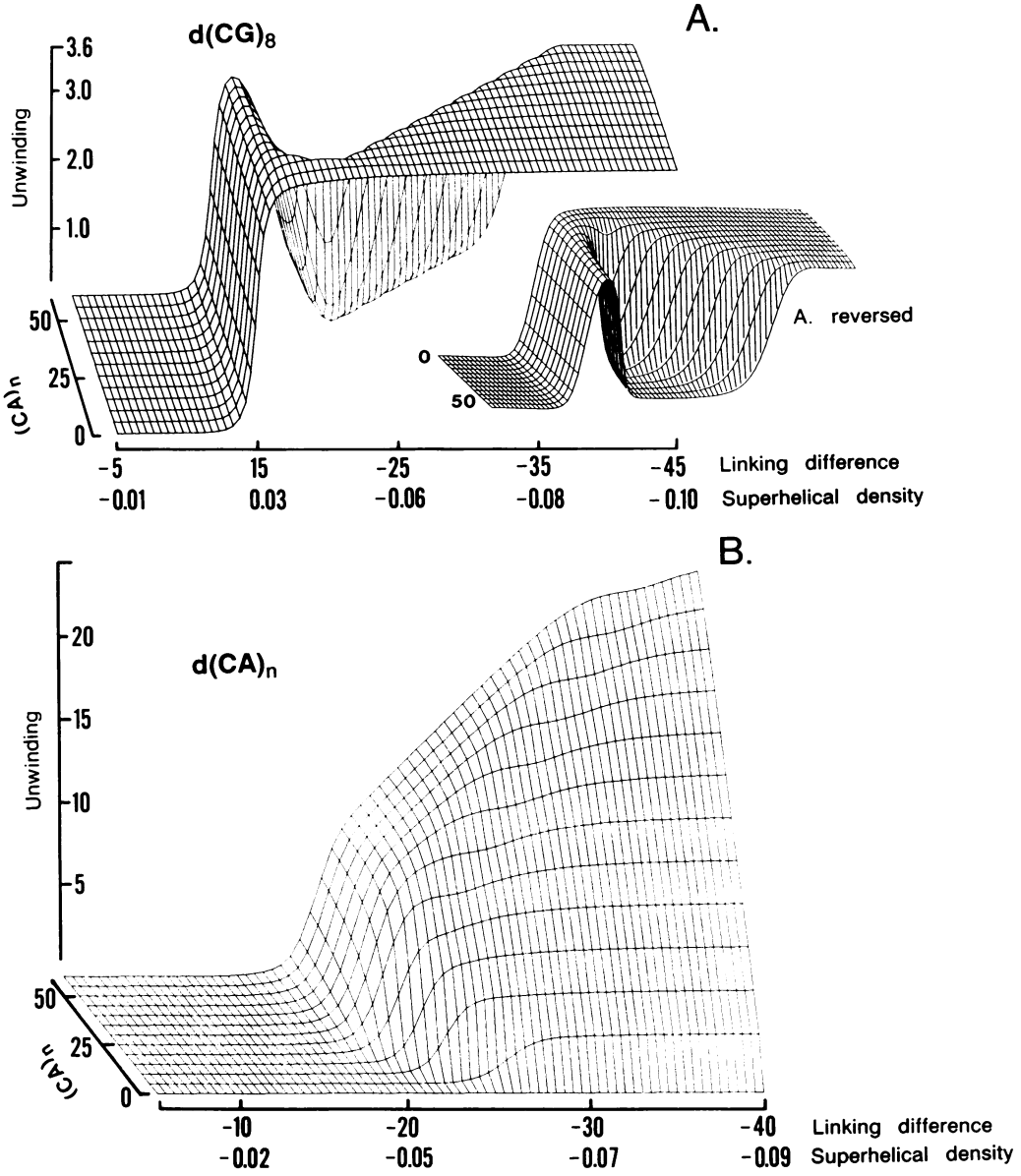

Fig. 1. Shown are the predicted behaviors of $(CG)_8$ versus $(CA)_n$ as a function of negative superhelicity when both sequences are found situated non-contiguously within the context of a 4.3-kb plasmid. Curves were calculated using a FORTRAN version of the statistical mechanical model described previously (Kelleher *et al.*, 1986). **(A)** The extent of unwinding due to Z-DNA formation in $(CG)_8$ is shown for varying lengths of $(CA)_n$. The inset shows the reverse side of curve A. **(B)** The extent of Z-DNA formation in the $(CA)_n$ sequence is shown for varying values of n. In each case Z-DNA formation is expressed as the number of turns of unwinding occurring in each sequence. The parameters used for these predictions are as follows: for $d(CG)_n$ (Peck and Wang, 1983) ΔG_{B-Z} = 0.33 kcal/mol bp, ΔG_j = 5.0 kcal/mol junction, a = 0.18 turn/bp (the unwinding associated with each base pair in the Z form relative to the B form) and b = 0.4 turn/junction (the unwinding associated with each B−Z junction). For $(CA)_n$ (Vologodskii and Frank-Kamenetskii, 1984), ΔG_{B-Z} = 0.67 kcal/mol bp, ΔG_j = 5.0 kcal/mol junction. Parameters a and b are assumed to be the same as for $(CG)_n$.

supercoiling, the property which determines gel mobility. For any given topoisomer, the reduction in twist resulting from the formation of Z-DNA produces a corresponding decrease in the number of negative supercoils. The formation of Z-DNA in one or both segments will therefore retard the mobility of topoisomers in the first dimension relative to topoisomers with linking differences insufficient to induce the formation of Z-DNA. The mobility and linking difference associated with each topoisomer can be unambiguously enumerated by electrophoresis in a second dimension under conditions in which Z-DNA is no longer favored such as in the presence of the intercalator chloroquin. The utility of the two-dimensional gel approach to an investigation of multiple transitions has been noted previously (Zacharias *et al.*, 1984; Kelleher *et al.*, 1986).

Comparisons between the unwinding data derived from two-dimensional gels and the unwinding of each sequence predicted by the model are restricted by two considerations. First, the gel conditions used to separate topoisomers of pZ2a and pZ2b in the first dimension must be identical to those used previously to determine the component energetic parameters used in the model. In addition, such gels are incapable of resolving the separate contributions to changes in twist arising from Z-DNA formation in each of the two sequences. Instead the mobilities of topoisomers undergoing the B−Z transition reflect the sum

of Z-DNA formation from both sequences. Despite this final limitation, it will become apparent from the arguments presented below that information regarding the experimentally observed behavior of each Z-forming sequence is nonetheless attainable.

Shown in Figure 3 are examples of two-dimensional electrophoretic gels of the topoisomers from plasmids pZ2a and pZ2b. It is apparent from these gels that the pattern of topoisomer mobility of plasmids pZ2a and pZ2b is considerably more complex than the distribution of topoisomers from the control plasmid pBR322 (Figure 3A and B). Plasmid pZ2a, for example, displays seven shifts in topoisomer mobility between linking differences of −14 to −32. Each of these trends in migration has been assigned the letters a−g. The patterns of topoisomer migration exhibited by pZ2a and pZ2b are virtually identical up to a linking difference of −21 turns. Beyond this value, the region of constant topoisomer migration labeled d, is considerably extended for pZ2b compared to pZ2a. Topoisomers are unresolved for pZ2b beyond this region. However, when conditions are selected that resolve topoisomers with greater negative linking differences (Figure 3C) additional transitions in migration similar to those observed at positions e, f and g for pZ2a are also observed for pZ2b. The degree of unwinding resulting from Z-DNA formation at each topoisomer is calculated by comparing the migration of pZ2a and pZ2b topoisomers to the migration of pBR322 topoisomers of similar linking difference (discussed in Materials and methods).

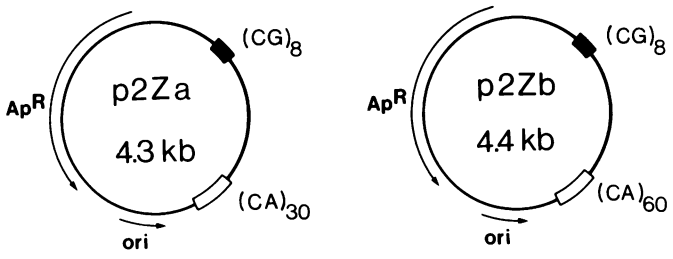

Fig. 2. Plasmids containing two non-contiguous Z-forming segments. Ap^R and ori refer respectively to the β-lactamase gene and replication origin derived from the plasmid pBR322. The Z-forming inserts are separated by ~600 bp.

Previous studies have characterized the Z-forming behavior of each of the sequences examined here when they are contained uniquely within plasmids (Haniford and Pulleyblank, 1983b; Peck and Wang, 1983; Vologodskii and Frank-Kamenetskii, 1984). Based on this previous work it is possible to assign qualitatively the patterns in topoisomer migration observed in Figure 3 to specific B−Z transitional events. Region a, for instance, corresponds to the formation of Z-DNA in the $d(CG)_8$ segment of pZ2a and pZ2b. Region c, on the other hand, can be ascribed to the formation of Z-DNA in the $(CA)_n$ segment. The regions of increasing topoisomer mobility, b and g, indicate that Z formation is essentially complete in $d(CG)_8$ and $d(CA)_n$ respectively. The migrational inflections denoted as e and f, however, are unique to the coupled B−Z equilibria.

Shown in Figure 4 are comparisons between the experimentally derived unwinding of plasmids pZ2a and pZ2b as a function of the linking difference and the theoretically predicted behavior of the coupled Z-DNA segments for each plasmid. In accord with the model shown in Figure 1, the predicted formation of Z-DNA in $d(CG)_8$ exhibits a trough initiating for both plasmids at a linking difference of −20 turns. The central difference between pZ2a and pZ2b as predicted by the model is that this trough extends over a greater range of topoisomers for pZ2b than for pZ2a. In comparison Z-DNA formation in the $d(CA)_n$ segments of both plasmids consistently increases as a function of linking difference.

We have compared the experimental unwinding data for each plasmid and the predicted combined unwinding from each segment as a function of the linking difference (produced by summing the individual curves for each segment shown in Figure 4B). The resulting curve representing the total theoretical unwinding for each plasmid is triphasic. A comparison of the ex-

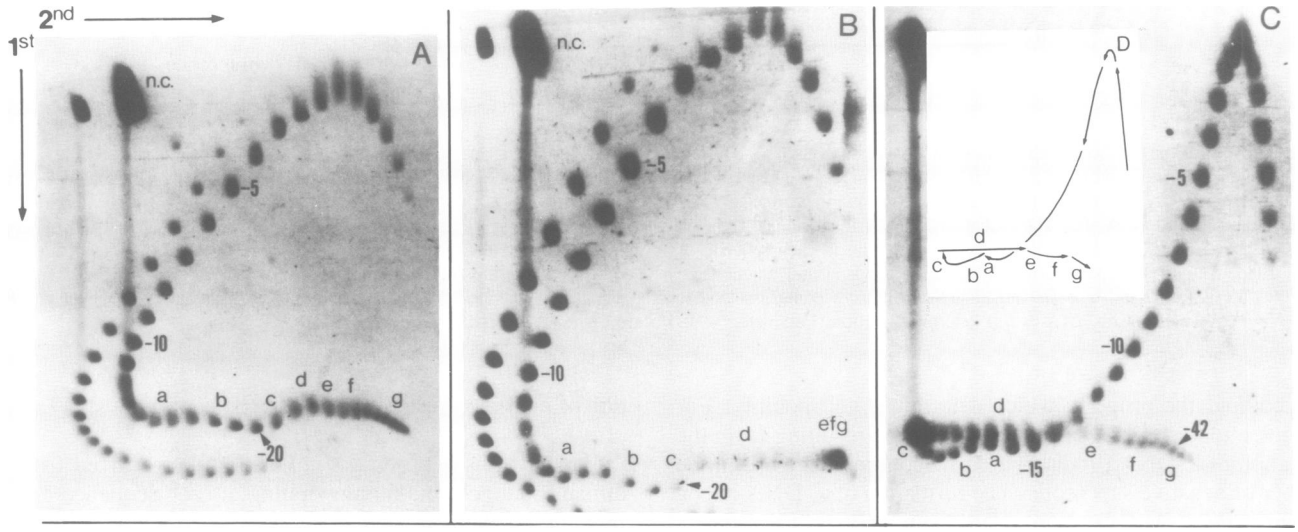

Fig. 3. Two-dimensional gel electrophoresis of plasmid topoisomers showing the coupled B−Z transitions of $(CG)_8$ versus $(CA)_{30}$ (plasmid pZ2a; gel photograph **A**) and $(CG)_8$ versus $(CA)_{60}$ (plasmid pZ2b; gel photographs **B** and **C**). The approximate linking difference from the relaxed configuration (topoisomer apex) has been indicated for selected topoisomers on each gel. 'n.c.' denotes the position of nicked circular DNA. Letters a−g correspond to trends in the rate of topoisomer migration. Also present on the left portion of gels **A** and **B** are topoisomers from pBR322. Topoisomers from the relaxed position of this distribution were intentionally omitted to prevent overlap with the apexes of pZ2a and pZ2b. In gels **A** and **B** the second dimension was electrophoresed in the presence of 0.88 μg/ml chloroquin. The second dimension of gel **C** was electrophoresed in the presence of 4.8 μg/ml chloroquin to resolve topoisomers with higher negative supercoiling. The pattern of topoisomer migration under these conditions has been simplified schematically in insert **D**. At position c in gel **C** several spots are superimposed at the relaxed position of the second dimension. Thus the label −42 is only approximate.

perimental data and these curves shows that there is a direct correspondence between the migrational transitions previously noted on the two-dimensional gels and the trends in slope evident from the theoretical predictions. Particular notice should be taken of the fact that the migrational transitions denoted as e and f on gels of pZ2a and pZ2b are predicted to occur by the model, which specifies not only their relative magnitudes but their approximate position with respect to the linking difference. Based on the above comparisons it is concluded that Z-DNA formation in $d(CG)_8$ does indeed undergo complex modulation as a function of linking difference. Thus, the model proposed for the coupled equilibria of discrete, multiple transitions (Kelleher *et al.*, 1986) appears to represent a reasonably accurate description of actual plasmid behavior.

Sensitivity of $d(CG)_8$ to BssHII cleavage. The unconventional behavior of $d(CG)_8$ in the coupled situation can be tested by measuring the sensitivity to cleavage of the topoisomer mixtures with the restriction endonuclease, *Bss*HII (Azorin *et al.*, 1984). This enzyme cleaves the sequence 5′-dGCGCGC only when it exists in the right-handed helical form. Therefore only those topoisomers in which the $d(CG)_8$ segment adopts the Z form are expected to resist cleavage. From Figure 4 it can be predicted that topoisomers occupying only the regions b and g will be resistant to cleavage.

Figure 5 shows two-dimensional gels of topoisomers from pZ2b treated or untreated with *Bss*HII prior to electrophoresis. In contrast to the conditions employed above, the first dimension was electrophoresed in the presence of 1 mM $MgCl_2$ in order to match the conditions used for the enzymatic digestion. As in Figure 3, however, the second dimension of one of the gels (gel B) was electrophoresed in the presence of a greater concentration of chloroquin to resolve those topoisomers with large negative linking differences. While the topoisomer patterns arising from the undigested samples on each gel are generally similar to those displayed in Figure 3 (B and C), they differ in subtle respects from previous gels run in the absence of $MgCl_2$. These deviations originate from the effect of Mg^{2+} on both the superhelicity of the topoisomers and on the energetics of the coupled equilibria; they are dealt with in greater detail below.

Notwithstanding the effect that Mg^{2+} has on the coupled transitions, it is apparent from Figure 5 that the topoisomers most resistant to cleavage by *Bss*HII are those located in areas of the gel most closely related to regions b and g in Figures 3 and 4, as predicted.

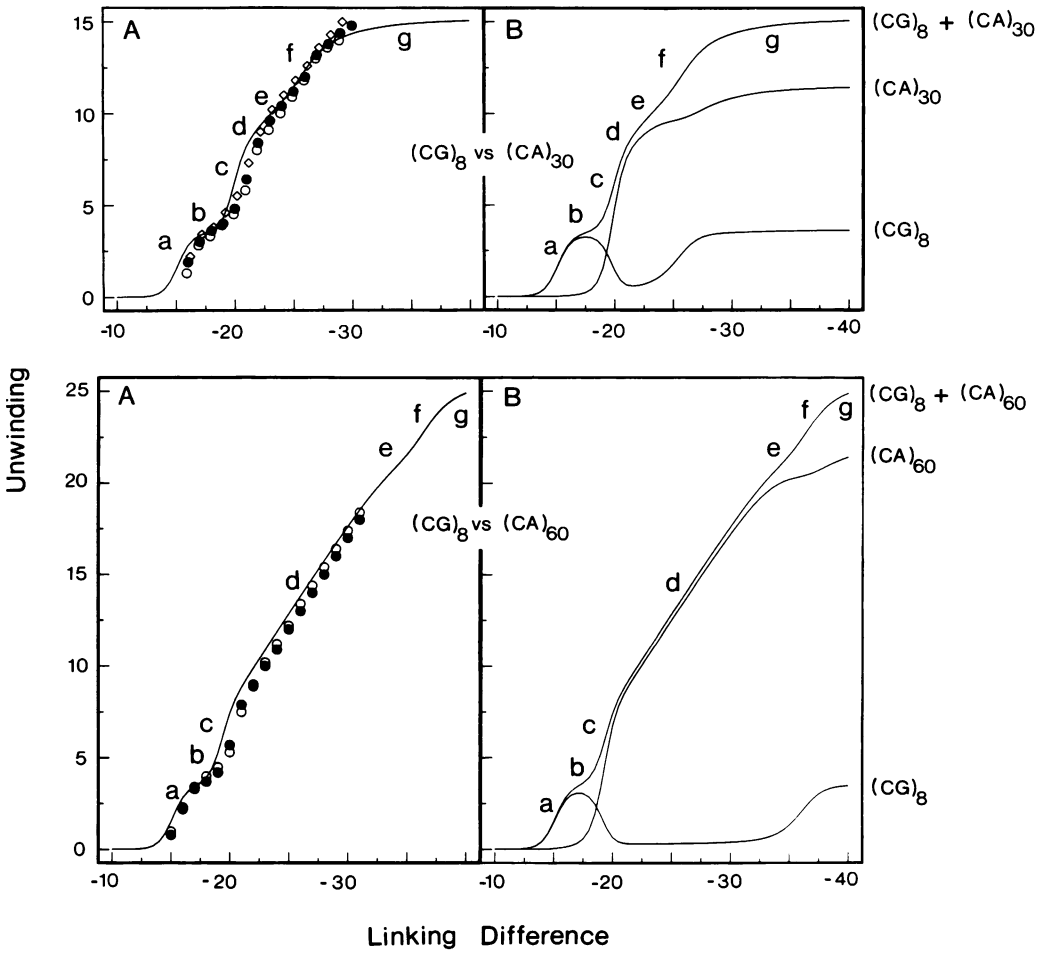

Fig. 4. Predicted and experimentally observed behavior of the coupled B−Z transition in plasmids pZ2a and pZ2b as a function of negative superhelicity. Unwinding data accumulated from two-dimensional gels for pZ2a **(A upper)** and pZ2b **(A lower)** are seen here plotted as a function of the negative linking difference of each plasmid. Each symbol type represents data obtained from a single gel. Filled circles correspond to data taken from the gels shown in Figure 2(A,B). Letters a−g refer to the mobility landmarks noted in Figure 3. Also shown in **A** is the predicted total unwinding arising from Z-DNA formation in both inserts as a function of negative superhelicity. Shown in **B** is the deconvolution of the total unwinding observed in **A** into the unwinding contributions of the $(CG)_8$ segment and $(CA)_n$ segment of each plasmid. Predicted curves were calculated as described in Figure 1.

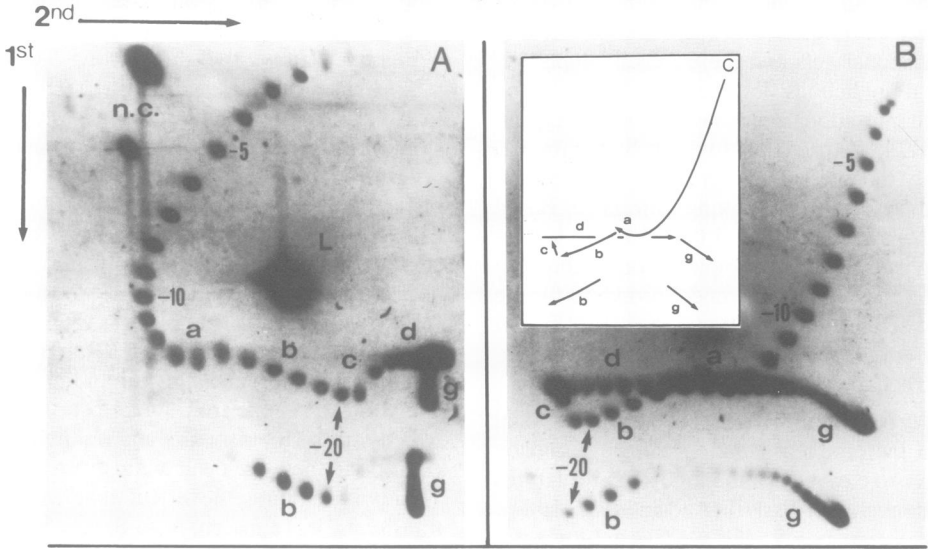

Fig. 5. Sensitivity of topoisomers to cleavage by *Bss*HII. Topoisomer mixtures of pZ2b were digested with *Bss*HII in the presence of 1 mM $MgCl_2$ as described in Materials and methods. Digested and undigested samples were electrophoresed on the same gel. The first dimension was run in the presence of 1 mM $MgCl_2$ as in Figure 3. The second dimension was electrophoresed in the presence of 0.2 μg/ml chloroquin **(A)**, or 3.8 μg/ml chloroquin **(B)**. Digested samples were loaded somewhat lower and slightly to the left of undigested samples. The assignment of linking numbers to topoisomers of the digested sample was made on the basis of this loading position. a–g correspond to the regions of transition shown in Figure 3. 'L' refers to linear DNA arising from *Bss*HII cleavage. The inset **C** indicates the direction of topoisomer migration with increasing negative superhelicity.

The effect of $MgCl_2$ on the coupled transitions. A close examination of topoisomer patterns from pZ2b run in the presence (Figure 5), or absence of $MgCl_2$ (Figure 3) reveals several significant differences in plasmid behavior between these two situations. The presence of $MgCl_2$, for instance, shifts the superhelical density of all topoisomers in the negative direction as evidenced by the loss of the positively supercoiled species located to the right of the apex of the topoisomer distribution. This phenomenon has previously been attributed to a slight overwinding of the normal B helix which produces a corresponding negative change in the degree of supercoiling (Shure and Vinograd, 1976).

In addition to this general effect, earlier studies have demonstrated that Mg^{2+} stabilizes the B–Z transition in polymers of $d(CG)_m$ (Behe and Felsenfeld, 1981; Van de Sande and Jovin, 1982). The behavior of $d(CA)_n$ in the presence of Mg^{2+} so far has not been reported. The effects of $MgCl_2$ on the coupled transition described here are readily apparent from the extension of region b in Figure 5 relative to the same region in Figure 3 and from the disappearance of the migrational inflections denoted as e and f. Not unexpectedly, the initial formation of Z-DNA in pZ2b occurs at slightly lower negative linking differences when $MgCl_2$ is present.

A major advantage of the model employed here to predict the competitive nature of supercoiling-induced transitions is that the free energy parameters used for each competing sequence were empirically derived and can be re-evaluated for various conditions. It was of considerable interest, therefore, to determine whether or not the discrepancies observed between the coupled transitions in the presence or absence of $MgCl_2$ could be accounted for by corresponding changes in ΔG_j and ΔG_{B-Z} for each sequence. Estimates of these two parameters for both $d(CG)_m$ and $d(CA)_n$ in the presence of 1 mM $MgCl_2$ were obtained from electrophoretic analysis of topoisomers from plasmids containing single inserts of either $d(CG)_m$ or $d(CA)_n$ (unpublished). For $d(CG)_m$, ΔG_j was found to shift unfavorably from 5 kcal/mol junction in the presence of $MgCl_2$ to ~6 kcal/mol junction in the presence of $MgCl_2$. Conversely, ΔG_{B-Z} shifted favorably from 0.33 kcal/mol bp to ~0.2 kcal/mol bp. $MgCl_2$ was found to have little effect on the energetics of $d(CA)_n$ and we have assumed here that the values associated with ΔG_j and ΔG_{B-Z} for this sequence remain unchanged from those specified in Figure 1.

By substituting these new energetic parameters for those employed previously in the model it is possible to predict the behavior of the Z-forming sequences contained in pZ2a and pZ2b in the presence of $MgCl_2$. In Figure 6, these results are compared with experimentally derived data for each plasmid in a manner identical to Figure 4. From the predicted case it can be seen that $d(CG)_8$ has become a more effective competitor for the free energy of supercoiling than in earlier examples. Z-DNA formation is detectable in this segment at lower values of negative linking difference and predominates in the Z form over a greater range of topoisomers before the equilibrium shifts in favor of $d(CA)_n$. Both these features are consistent with the experimentally observed extension of region b in Figure 5 relative to the same region in Figure 3. The enhanced competitiveness of $d(CG)_8$ in the theoretical case, however, is best illustrated by the lessened ability of the $d(CA)_n$ segment to shift the equilibrium away from Z-DNA formation in $d(CG)_8$. This effect is most marked in pZ2a and correlates well with the observation that topoisomers located within this region of linking difference show only slight *Bss*HII cleavage (results not shown). Even in pZ2b it is apparent that topoisomers situated in the extended region d show some resistance to cleavage relative to those topoisomers situated prior to the onset of transition. Additional validation of the model is provided by the predicted disappearance of the inflections e and f which is consistent with the absence of these migrational variations in the gels displayed in Figure 5.

Predicting the behavior of three Z-forming sequences. The model outlined in previous sections can be extended to a treatment of any number of transitions competing for the free energy of super-

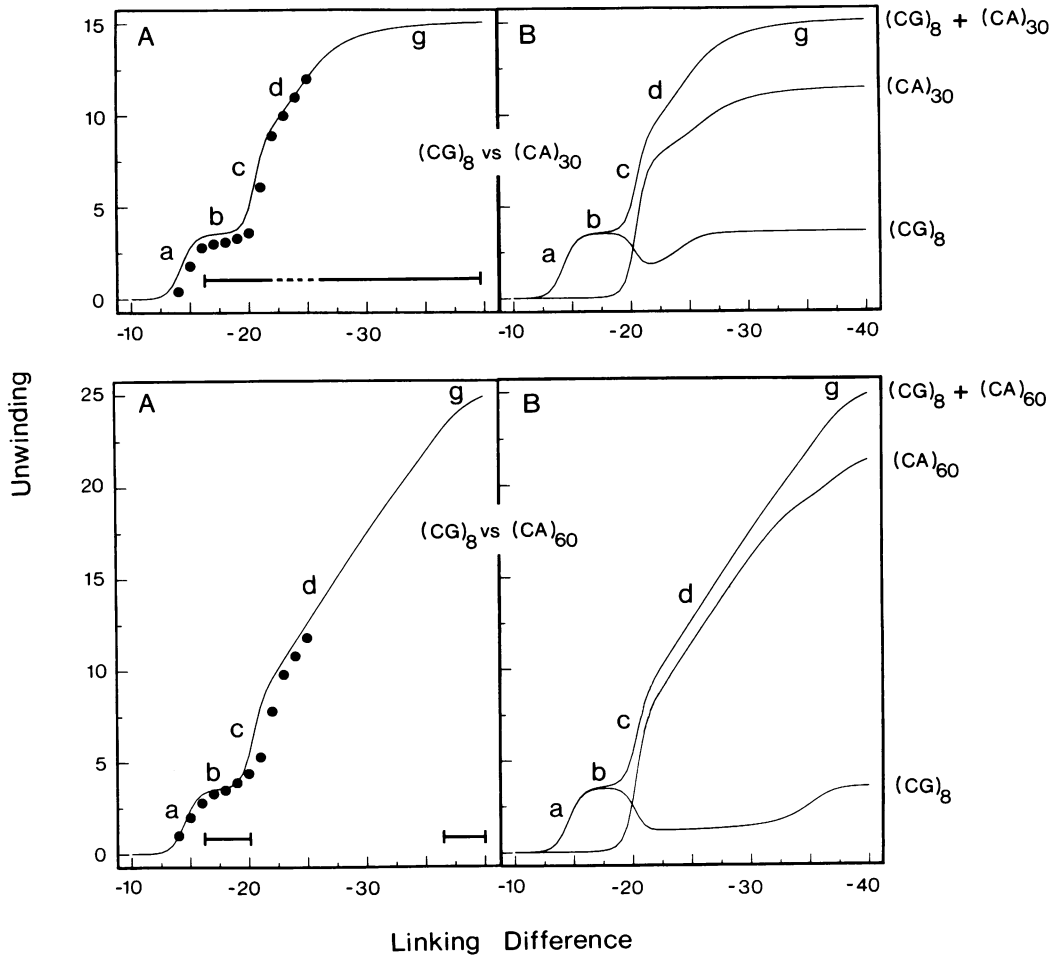

Fig. 6. The effect of $MgCl_2$ on the coupled transitions $(CG)_8$ versus $(CA)_{30}$ and $(CG)_8$ versus $(CA)_{60}$. Shown in **A** is the unwinding data accumulated from two-dimensional gels of topoisomers from pZ2a **(upper)** and pZ2b **(lower)**. Gel running conditions were identical to those used to collect the data shown in Figure 4 with exception that 1 mM $MgCl_2$ was present in the running buffer. Letters a–g are analogous to the mobility transitions shown in Figures 3 and 4 with the notable absence of the inflections denoted as e and f in previous figures. Predicted transition curves were calculated as in Figure 1. For these calculations, free energy parameters for the $(CG)_8$ segment were changed to reflect the altered electrophoresis conditions: ΔG_i = 6.0 kcal/mol bp. ΔG_{B-Z} = 0.2 kcal/mol bp. Bars at the base of **A** (upper) and **A** (lower) coincide with topoisomers found to be resistant to cleavage by *Bss*HII. The dashed region in the upper bar represents slight sensitivity to *Bss*HII cleavage.

coiling at different locations within the same topological domain. The only requirement for such an extension is a prior knowledge of the energetics associated with each of the sequences that comprise the system. The example in Figure 7 shows the predicted behavior of the sequences $d(CG)_8$, $d(CACG)_7$ and $d(CA)_{30}$ which have been placed non-contiguously within the context of a 4.3-kb plasmid.

The Z-forming behavior of $d(CG)_m$ and $d(CA)_n$ as a function of negative superhelicity is well understood both for the coupled case (this work) and the uncoupled case (Peck and Wang, 1983; Vologodskii and Frank-Kamenetskii, 1984). Information concerning the energetics of the B–Z transition in $d(CACG)_n$ is currently unavailable. It can be seen however that this sequence contains equal numbers of the nearest-neighbor interactions that comprise $d(CG)_m$ or $d(CA)_n$. We have assumed here for the sake of demonstration that the energetics associated with the transition of $d(CACG)_n$ are intermediate to those of $d(CG)_m$ and $d(CA)_n$ per unit length of sequence. In comparison to the two-sequence coupled system examined previously, the three-sequence case is considerably more complicated. Z-DNA formation in $d(CG)_8$ and $d(CACG)_7$ as a function of the linking difference gives rise to three local maxima in the first sequence and two in the latter sequence. The curve produced by the combined unwinding of all three sequences shows a reasonably unperturbed ascent with only minor inflections; thus the total unwinding occurring in the plasmid is buffered to some extent from the pronounced changes occurring within each segment.

Discussion

The results of the present work theoretically and experimentally illustrate that a supercoiling-dependent transition occurring in one sequence can act at a distance in a complex way to affect the behavior of other transitions occurring elsewhere within a given topological domain. This long-range effect arises not from the direct communication of structural information between sequences, but rather from the competition of discrete transitions for the free energy associated with negative supercoiling. With this in mind, such interactions are in principal unlimited by the distance separating the sites of local transition, or the relative orientation of the competing elements.

The coupled behavior of non-contiguous structural transitions can be understood within the context of equilibrium statistical mechanics. As demonstrated in the present work such an analysis

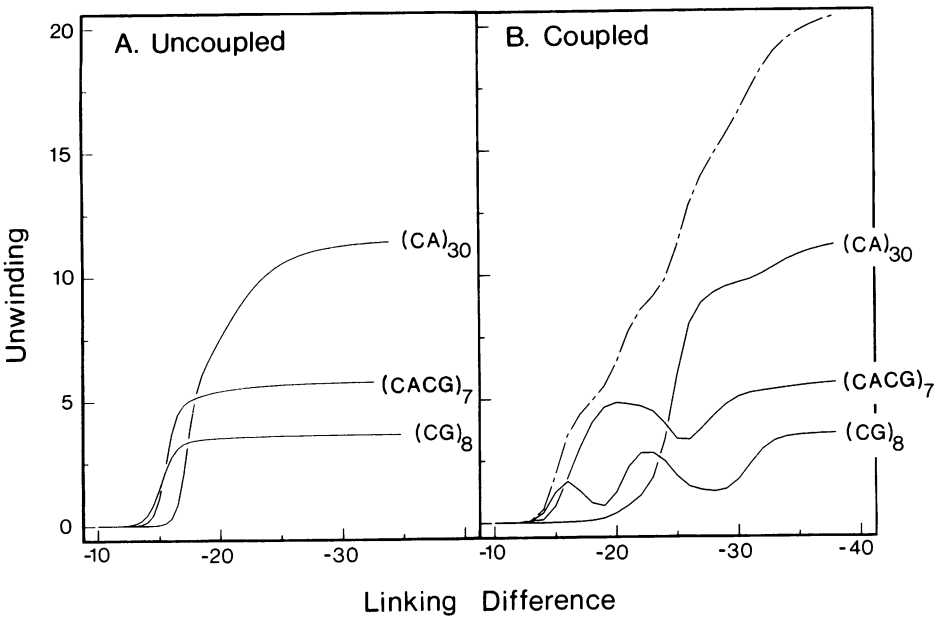

Fig. 7. The coupled B−Z transition of three different Z-DNA-forming sequences. Transitions have been modeled within the context of a 4363-bp plasmid. **A** shows the expected behavior of each sequence existing by itself in the plasmid. **B** shows the expected behavior of each sequence contained within the same plasmid. The broken line in **B** represents the sum of the individual contributions to the unwinding occurring in each sequence. The energetic parameters used for $(CG)_8$ and $(CA)_{30}$ are the same as those in Figure 1. In this case $(CACG)_7$ is presumed to consist of equal numbers of (CA) and (CG) nearest-neighbor interactions.

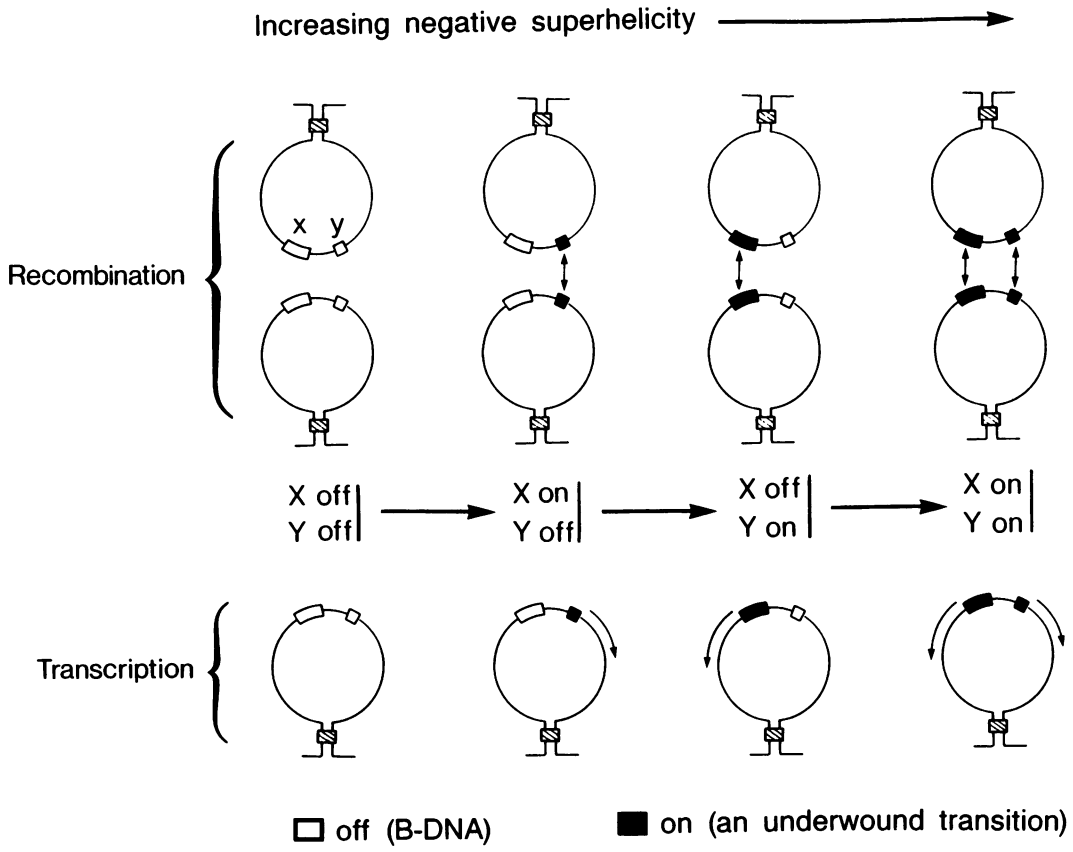

Fig. 8. A model for the involvement of topologically coupled transitions in the programmed switching of genetic events. Shown in columns 1−4 is a topologically discrete chromosomal domain containing two sequences X (small box) and Y (large box) that are capable of undergoing transition to an underwound form. The hatched box represents the anchor point of the topological domain. The negative superhelicity of this domain increases as shown. Sequences having undergone transition are shaded black. For the recombination case, arrows mark the sites of potential interaction between two homologous domains. For the transcriptional case, arrows indicate the onset of transcription assuming that X and Y are control sequences which activate the gene upon transition. The transcriptional directions indicated are arbitrary.

can provide a reasonably accurate description of competing B−Z transitions occurring in sequences that have been previously characterized thermodynamically. The nature of this competition as seen in Figure 1 is critically dependent on both the composition and length of each sequence contained in the coupled system. In the competition illustrated for $d(CG)_8$ versus $d(CA)_n$ for example, the behavior of $d(CG)_8$ is most complex at longer lengths of $d(CA)_n$. Such behavior can be rationalized in qualitative terms. The energetic requirements of $d(CG)_8$ for Z-DNA formation are less than those of $d(CA)_n$, therefore Z-DNA formation predominates in $d(CG)_8$ at lower values of the negative linking difference. The extent of torsional relief that can be afforded to supercoiled DNA by $d(CG)_8$, however, is limited to ~3.6 turns. Long stretches of $d(CA)_n$ on the other hand have a greater capacity to absorb the unfavorable free energy associated with negative supercoiling and so become effective competitors for the free energy of supercoiling at greater values of the negative linking difference. The maintenance of four B−Z junctions (two junctions per transition) imposes a large energetic penalty. As a result the B−Z equilibrium shifts to propagating Z-DNA in the longer $d(CA)_n$ segment at the expense of Z-DNA formation in $d(CG)_8$. Such competitive behavior will be characteristic of any topological system in which short sequences with strong transitional potential compete against longer sequences with a lower propensity for transition.

In general, changes in ambient conditions which alter the energetics of the coupled system will by necessity alter the behavior of transitions occurring within the system. As we have seen here, the presence of $MgCl_2$ has two effects on the coupled B−Z equilibria. First, $MgCl_2$ alters the level of negative supercoiling thus producing corresponding changes in ΔG_τ. In addition, $MgCl_2$ has a direct differential effect on the energetics of $d(CG)_8$ and $d(CA)_n$. Such alterations in conditions are easily accommodated provided that each of the relevant energetic parameters have been re-evaluated under the new conditions.

The present work focuses on a description of the competition between discrete sequences undergoing the B−Z transition. We emphasize, however, that the complicated behavior seen here is possible for any combination of structural transitions which alter the helical repeat of DNA and are therefore supercoiling dependent. Two other known examples of supercoiling-induced transitions include cruciform extrusion and an incompletely characterized transition in homo purine pyrimidine sequences (reviewed in Pulleyblank *et al.*, 1985). In addition, the early stages of strand separation events that arise from processes such as replication, recombination and transcription may also be included in this list. In so far as the competing transitions in question can be described by a two-state process, the model applied here is extendable to a variety of competitive situations.

A model for programmable genetic switching

The observed behavior of competing transitions as a function of negative superhelicity suggests a way in which discrete genetic events can be programmed to follow a defined and potentially complicated sequential pattern. The genomes of both prokaryotic and eukaryotic organisms are believed to be organized into topologically discrete loops ranging in size from 50 to 100 kb (reviewed in Gasser and Laemmli, 1986). Let us now consider a situation where such a loop contains two discrete sequences X and Y, each capable of transition to an underwound form. Since X and Y are coupled only through topology, they can be situated with equal effect in any orientation or at any distance with respect to each other.

In the simplest case we can imagine that the transition of each sequence to their underwound forms is an all or none process where the B-DNA form of each sequence is the 'off' position and the underwound structure is the 'on' position. Within this binary system there exist four potential coupled states for sequences X and Y: X off/Y off; X on/Y on; X off/Y on; and X on/Y off. The program that determines which of these four coupled states are to be permitted, and in what order they appear as a function of superhelicity, is determined by the energetic requirements of each sequence and the magnitude of unwinding associated with each transition. The program itself is set in motion by perturbing the superhelical density of the loop from a pre-defined value in the positive or negative direction. The superhelical density could be varied by topoisomers which alter the linking number or by altering the helical repeat of B-DNA at a constant value of linkage. Alternatively, the progressive assembly or disassembly of nucleosomes could also vary the torsion and bending strain in protein-free regions. The energetic requirements of each sequence *in vivo* will be considerably more complicated than those previously discussed for the *in vitro* case. Not only will they depend on sequence length, base composition and the ambient conditions which surround each sequence, but they will also depend upon the effects of base modification and protein stabilization of one structural form over another.

A simple illustrative example of one type of switching program is shown in Figure 8 where sequences X and Y exhibit behaviors that are similar to $(CG)_8$ and $(CA)_n$ respectively. In this particular case the program order takes the form X off/Y off to X on/Y off to X off/Y on to X on/Y on, as the superhelical density is increased. Thus all four coupled states are represented. As we have seen from the three-transition case displayed in Figure 7, such program can take on a complex character as competing sequences are added to the system. Moreover, it can be appreciated that any change in the energetic requirements of a single sequence, whether it be through mutation, base modification or protein binding, is likely to affect the pattern of events in all other sequences competing within a given topological domain.

Recent observations have made it possible to apply this model to the process of generalized genetic recombination. Rec1, a DNA recombination protein isolated from the lower eukaryote *Ustilago maydis*, has been demonstrated to bind to Z-DNA in an ATP-dependent manner (Kmiec *et al.*, 1985). Moreover, Rec1 has been shown to direct the bimolecular association of negatively supercoiled DNA molecules *in vitro* to homologous sites which show purine−pyrimidine alternation (Kmiec and Holloman, 1986). Whether Z-DNA must exist in each molecule prior to recombination or exists as a stable intermediate of recombination remains to be clarified.

Regardless of the precise mechanism by which Rec1 functions it is clear that the initiation of recombination requires that both parental molecules be underwound at the site of interaction. Within a given topologically closed domain there could exist several favored sites of recombination which will compete to some extent with each other for the available free energy of supercoiling. Based upon the above considerations it can be envisioned that recombination will be favored at one site at a particular value of the superhelical density and at other sites as the superhelical density is varied. In this model the location of genetic recombination is determined solely on the basis of topological considerations.

Naturally, the above ideas can also be applied to the coordinated regulation of gene expression for those situations where transcription is sensitive to changes in the topological state of the DNA.

In this event, transitions occurring within regulatory sequences may create or destroy the recognition sites of the proteins necessary for specific gene function. In Figure 8, for example, the sequences X and Y are indicated to serve as the control regions for two hypothetical genes which become activated upon transition to their respective alternative structural forms.

In applying the competitive model to transcriptional and recombinational processes we have made a number of suppositions which relate to the states of DNA within a cell. We assume that structural transitions are permitted to occur *in vivo* and that the superhelical density of topological domains can be varied over a range that makes full use of their competitive nature. We also assume that the rate at which conformational changes occur in DNA is rapid when compared to the rate of topoisomerase action. To date the validity of these assumptions remains uncertain, therefore the extent to which coupled transitional equilibria exist *in vivo* and co-ordinate genetic processes is presently unknown.

Materials and methods

Plasmid constructions

pZ2a and pZ2b were constructed by removing the 622-bp fragment situated between the *Ball* site and the *Pvu*II site of pLP316 and replacing it with either the 567-bp *Hinc*II fragment from pDHf14 (Haniford and Pulleyblank, 1983b) to produce pZ2a, or the 627-bp *Hinc*II fragment from pDHf8 (Haniford and Pulleyblank, 1983b) to produce pZ2b. Plasmid pLP316 is identical to pLP332 (Peck and Wang, 1985) with the exception of the length of the $(CG)_n$ insert. The insert orientation of pZ2a and pZ2b is such that the $(CA)_n$ segment of each plasmid is proximal to the *Pvu*II site.

Preparation of plasmid topoisomers

A population of topoisomers was prepared for each plasmid by adding varying amounts of ethidium bromide to aliquots of plasmid in the presence of topoisomerase I (Bethesda Research Labs). Typically, 2 units of topoisomerase I were added to 5 μg of plasmid in 50 μl of reaction buffer (0.2 M NaCl, 10 mM Tris, 1 mM EDTA, pH 8.0). Ethidium bromide was added in amounts which varied from 0 to 25 μM final concentration in 1 μM increments. Reactions were performed at room temperature overnight in the dark. The reactions were terminated by phenol extraction, and plasmid aliquots were pooled and extracted with butanol to remove the ethidium bromide. The DNA was then precipitated with ethanol and resuspended to a final concentration of 500 μg/ml in 10 mM Tris, 1 mM EDTA, pH 8.0.

BssHII digestion of topoisomers

Ten micrograms of plasmid topoisomers prepared as described above were digested for 2 h at room temperature with 40 units of *Bss*HII (New England Biolabs) in 0.3 ml total volume of TBM buffer (90 mM Tris borate, 1 mM $MgCl_2$, pH 8.3).

Two-dimensional gel electrophoresis

Agarose gel electrophoresis was performed essentially as described previously (Ellison *et al.*, 1985) but with several notable exceptions. First dimensions in the presence of TBM buffer were electrophoresed for 40 h at 110 V in order to achieve comparable resolution to gels run for 22 h in TBE buffer (90 mM Tris borate, 2.5 mM EDTA, pH 8.3). Prior to running the second dimension, gels were equilibrated in 2500 ml of TBE buffer and varying amounts of chloroquin (see figure legends). In general, first dimensions run in the presence of $MgCl_2$ required less chloroquin in the second dimension to achieve comparable separation to gels electrophoresed in the absence of $MgCl_2$.

Experimental unwinding measurements

Unwinding resulting from Z-DNA formation was determined from electrophoretic gels by plotting the mobility of topoisomers for a given plasmid (standardized to the mobility of the nicked species) as a function of the linking difference, as previously described (Haniford and Pulleyblank, 1983b). The unwinding exhibited by a given topoisomer is obtained by subtracting its linking difference from the linking difference of a pBR322 topoisomer with an identical corrected mobility.

Theoretical calculations

The Model program used in the present work is a FORTRAN adaptation of the statistical mechanical model previously described (Kelleher *et al.*, 1986). In addition, the program was modified to deal with up to three competing transitions. The program time for the three-sequence situation displayed in Figure 7 was approximately 300 CPU h on a Vax 11-780. This program is available on request.

References

Azorin,F., Hahn,R. and Rich,A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5714–5718.

Baker,T.A., Sekimizu,K., Funnell,B.E. and Kornberg,A. (1986) *Cell*, **45**, 53–64.

Behe,M. and Felsenfeld,G. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 1619–1623.

Benham,C.J. (1981) *J. Mol. Biol.*, **150**, 43–68.

Benham,C.J. (1982) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 219–227.

Depew,R.E. and Wang,J.C. (1975) *Proc. Natl. Acad. Sci. USA*, **11**, 4275–4279.

Ellison,M.J., Kelleher,R.J.,III, Wang,A.H.-J., Habener,J.F. and Rich,A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 8320–8324.

Ellison,M.J., Feigon,J., Kelleher,R.J.,III, Wang,A.H.-J., Habener,J. and Rich,A. (1986) *Biochemistry*, **25**, 3648–3655.

Gasser,S.M. and Laemmli,U.K. (1986) *EMBO J.*, **5**, 511–518.

Gellert,M. (1981) *Annu. Rev. Biochem.*, **50**, 879–910.

Haniford,D.B. and Pulleyblank,D.E. (1983a) *J. Biomol. Struct. Dynam.*, **1**, 593–609.

Haniford,D.B. and Pulleyblank,D.E. (1983b) *Nature*, **302**, 632–634.

Haniford,D.B. and Pulleyblank,D.E. (1985) *Nucleic Acids Res.*, **13**, 4343–4363.

Kelleher,R.J.,III, Ellison,M.J., Ho,P.S. and Rich,A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 6342–6346.

Klysik,J., Stirdivant,S.M., Larson,J.E., Hart,P.A. and Wells,R.D. (1981) *Nature*, **290**, 672–676.

Kmiec,E.B. and Holloman,W.K. (1986) *Cell*, **44**, 545–554.

Kmiec,E.B., Angelides,K.J. and Holloman,W.K. (1985) *Cell*, **40**, 139–145.

Peck,L.J. and Wang,J.C. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6206–6210.

Peck,L.J. and Wang,J.C. (1985) *Cell*, **40**, 129–137.

Pulleyblank,D.E., Shure,M., Tank,D., Vinograd,J. and Vosberg,H.-P. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 4280–4284.

Pulleyblank,D.E., Haniford,D.B. and Morgan,A.R. (1985) *Cell*, **42**, 271–280.

Rich,A., Nordheim,A. and Wang,A.H.-J. (1984) *Annu. Rev. Biochem.*, **53**, 791–846.

Shure,M. and Vinograd,J. (1976) *Cell*, **8**, 215–226.

Singleton,C.K., Klysik,J., Stirdivant,S.M. and Wells,R.D. (1982) *Nature*, **299**, 312–316.

Sullivan,K.M. and Lilley,D.M.J. (1986) *Cell*, **47**, 817–827.

Van de Sande,J.H. and Jovin,T.M. (1982) *EMBO J.*, **1**, 115–120.

Vologodskii,A.V. and Frank-Kamenetskii,M.D. (1984) *J. Biomol. Struct. Dynam.*, **1**, 1325–1333.

Wang,J.C. (1982) *Unwinding at the Promoter and the Modulation of Transcription by DNA Supercoiling*. Praeger, New York.

Wang,J.C. (1985) *Annu. Rev. Biochem.*, **54**, 665–697.

Wang,A.H.-J., Quigley,G.J., Kolpak,F.J., Crawford,J.L., van Boom,J.H., van der Marel,G. and Rich,A. (1979) *Nature*, **282**, 680–686.

Weisberg,R. and Landy,A. (1983) In Stahl,F.W. and Weisberg,R.A. (eds), *Lambda II*. Cold Spring Harbor Laboratory Press, NY.

Zacharias,W., Larson,J.E., Kilpatrick,M.W. and Wells,R.D. (1984) *Nucleic Acids Res.*, **12**, 7677–7692.

Received on January 8, 1987; revised on February 16, 1987