CA Runs Increase DNA Flexibility in the Complex of λ Cro Protein with the O_R3 Site†

Yuri L. Lyubchenko,‡,§,∥,⊛ Lyuda S. Shlyakhtenko,‡,⊛ Ettore Appella,▽ and Rodney E. Harrington*,‡,§

Department of Biochemistry, University of Nevada–Reno, Reno, Nevada 89557, Departments of Microbiology and Physics, Arizona State University, Tempe, Arizona 85287, Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, and Institute of Molecular Genetics, Russian Academy of Sciences, Moscow

Received December 3, 1992; Revised Manuscript Received February 5, 1993

ABSTRACT: The alternating pyrimidine–purine elements CA, CAC, and CACA are anisotropically flexible, as deduced from gel circularization assays on point mutations and single-base mismatches in the O_R3 site of λ phage alone and in the specific complex with the Cro protein. These sequences evidently promote DNA bending in the specific binding region of the complex and may also facilitate overwinding in the central nonbinding region. Effects for CACA are exceptionally large and suggest that an alternative DNA structure may occur in this element.

Conformation and structural changes in DNA are known to accompany the specific binding of many proteins [reviewed in Travers (1989), Harrison and Aggarvwal (1990), and Steitz (1990)]. These changes may include axial bending and localized changes in helical twist angles. When they occur, they may improve the fit between direct recognition elements in the DNA and the protein, leading to stronger hydrogen-bonding, electrostatic, and van der Waals interactions (Harrison & Aggarvwal, 1990; Schultz et al., 1991; Cheung et al., (1984). Collectively, they have been termed "indirect recognition" (Drew & Travers, 1985), and they evidently comprise the basis for much of the specificity and selectivity in protein–DNA interactions. In spite of much recent progress in understanding indirect recognition, facilitated by direct structural studies (Steitz 1990) and high-resolution NMR (Cheung et al., 1984; Donlan & Lu, 1992; Leroy et al., 1988a,b; Kirpichnikov et al., 1964; Metzler & Lu, 1989), many aspects still remain unclear.

The analysis of ring closure, or cyclization, in DNA fragments containing several hundred base pairs (Shore & Baldwin, 1983a,b; Horowitz & Wang, 1984; Levene & Crothers, 1986a,b; Hagerman & Ramadevi, 1990; Taylor & Hagerman, 1990) has been shown to be a sensitive tool for the determination of both torsional and bending flexibility in DNA (Schellman, 1974; Zhurkin 1985). The basic methods were pioneered by Shore et al. (1983a,b). Variants and improvements have been described by Levene and Crothers (1986a,b) and by Hagerman et al. (Hagerman & Ramadevi, 1990; Taylor & Hagerman, 1990). An alternative approach was described by Ulanovsky et al. (1986), who determined the cyclization properties of ligated fragments having two integral helical repeats (21 bp) using two-dimensional polyacrylamide gel electrophoresis and used these to estimate fixed bending in poly(A) tracts. This approach was later used to study DNA bending at the λ replication origin (Zahn & Blattner, 1988) and at photoinduced thymine dimers (Hussain et al., 1988). Gel band shift measurements have also been used to characterize mismatches and looped regions (Hsien & Griffith, 1989; Bhattacharyya & Lilley, 1989). We recently employed a variant of the Ulanovsky et al. (1986) procedure to estimate the bending angle induced in the O_R3 site by the Cro protein of λ phage (Lyubchenko et al., 1991). In the present study, we show that this approach can also provide important information on sequence-specific flexibility in DNA.

Table I: Complete Set of Oligonucleotide Sequences Studied in This Work[a]

5'- TATCACCGCAAGGGATAAATA -3'	wild type
3'- TGGCGTTCCCTATTTATATAG -5'	
Mutations	
5'- TATCACCaCAAGGGATAAATA -3'	M1
5'- TATCACCGgAAGGGATAAATA -3'	M2
5'- TATCACCGCAAGtGATAAATA -3'	M3
5'- TATCtCCGgAAGGGATAAATA -3'	M4
Mismatches	
5'- TATCACCACAAGGGATAAATA -3' (C)	COM1
5'- TATCACCGGAAGGGATAAATA -3' (G)	COM2
5'- TATCACCGCAAGTGATAAATA -3' (C)	COM3

[a] The specific binding regions (Takeda et al., 1989; Brennan et al., 1990) are indicated by bold/underlined typeface. Mutation sites are shown in lowercase typeface. The complementary strand is shown only for the wild type (WT) sequence. Complementary strands for all others were constructed with identical overhanging ends.

METHODS

Design and Preparation of Oligonucleotides. The complete set of oligonucleotide sequences studied in this work are shown in Table I. All were synthesized on an Applied Biosystems 392 DNA synthesizer and purified on a Waters Model 510

† Supported, in part, by Hatch 126 from the Nevada Agricultural Experiment Station, by Research Grants GM 33435 from the National Institutes of Healt and DMB 9117488 from the National Science Foundation (R.E.H), and by Research grant N000149J1455 from Office of Naval Research (Stuart M. Lindsay, PI) in partial support of Y.L.L.

* Corresponding author.
‡ University of Nevada—Reno.
§ Department of Microbiology, Arizona State University.
∥ Department of Physics, Arizona State University.
⊛ Russian Academy of Sciences.
▽ National Institutes of Health.

0006-2960/93/0432-4121$04.00/0

HPLC. The specific binding regions (Takeda et al., 1989; Brennan et al., 1990) are indicated by bold underlined typeface. Mutation sites are shown in lowercase typeface. The complementary strand is shown only for the wild-type (WT) sequence. Complementary strands for all others were constructed with identical overhanging ends. Mutation M1 introduces a CACA/TGTG element beginning at position 7 in tandem with the wild-type CAC/GTG element beginning at position 4. Mutation M3 introduces a GTG/CAC element beginning at position 12. Mutation M2 eliminates the wild-type CA/GT at position 9, and mutation M4 eliminates all wild-type CA/GT and CAC/GTG elements. Mismatched double-stranded fragments COM1–COM3 were constructed by hybridizing the indicated single-stranded oligonucleotides from M1, M2, and M3 with complementary strands from the wild-type O_R3 sequence. The single base shown below each of these is one base of the complementary strand and indicates the specific mismatch site. Mismatch COM1 introduces a mismatch at position 8 in the CACA/TGTG element of M1; COM 2, a mismatch at position 9 of M2; and COM3, a mismatch at the center of the GTG/CAC element of M3 at position 13. Mutations M1 and M2 leave the binding constant essentially unchanged compared to the wild type, while M3 and M4, respectively, lead to a small increase and a small decrease as suggested by the data of Takeda et al. (1989) and confirmed by us in gel band shift assays (data not shown). These same gel assays show that Cro binds with similar efficiency to the mismatches COM1–COM3.

Ligation and Analysis by Two-Dimensional Gel Electrophoresis. Polyacrylamide gel cyclization assays followed the method of Ulanovsky et al. (1986) as modified by Lyubchenko et al. (1991). Ligation products of labeled 21-bp oligonucleotides containing 17-bp wild-type and mutated or mismatched O_R3 operator sequences (Table I) were analyzed in both the presence and the absence of bound Cro protein. These precursor oligonucleotides were labeled with ^{32}P at the 5′, ends with polynucleotide kinase, annealed with the complementary strands, and ligated at 0 °C for 12 h in the presence or absence of Cro protein. The ligation mixture of circles and linear molecules was separated first on 4% polyacrylamide gel (first dimension) and then on 10% polyacrylamide gel (second dimension). In the second dimension, circular molecules are well resolved both from each other and from linear species. The number of 21-bp elements (21-mers) in each oligomer species is indicated. Lengths of circular molecules were verified in all cases by running the DNA extracted from spots excised from these gels in denaturing gels. Additional details have been described elsewhere (Lyubchenko et al., 1991).

Cyclization efficiency for each circle (yield of circles) is determined from the radioactivity of DNA extracted from the corresponding spot excised from the gel normalized by the averaged total radioactivity of DNA from spots corresponding to 5-, 6-, and 7-mer linear oligomers. For comparison we also normalized the yield of each circle by the sum of the intensities of all remaining linear and circular species. The shape of the distribution was usually invariant regardless of the method of normalization, but in some cases it was difficult to evaluate the spot intensities of 11-, 12-, and 13-mers for linear molecules because of saturation. Total radioactivity was determined on a Beckman LS 1500 counter.

Circular Dichroism Studies. All CD measurements were obtained on a Jasco J600 spectropolarimeter equipped with a special microcell attachment. Spectra were obtained by scanning between 210 and 310 nm, but only the longer wavelength part of each spectrum, which reflects DNA

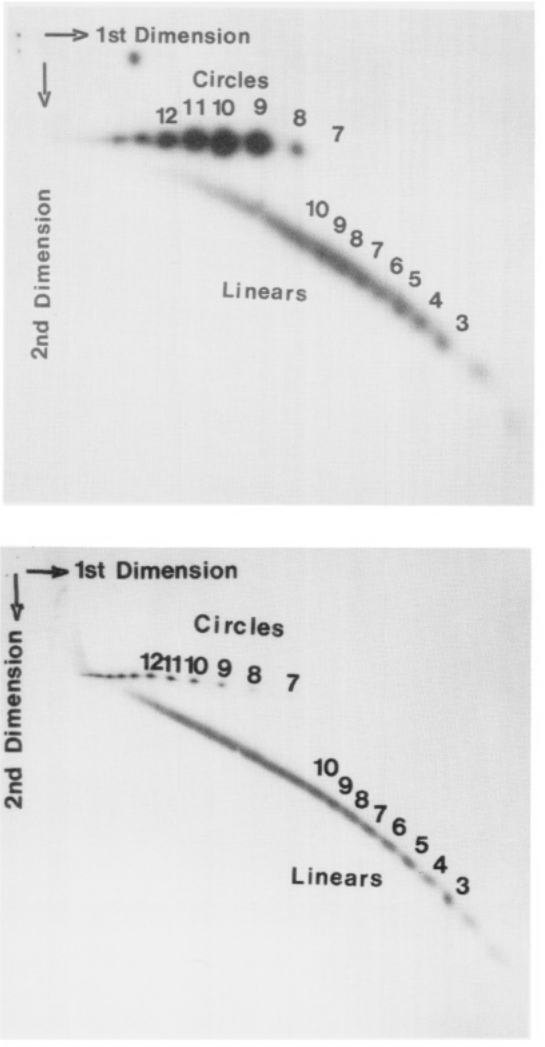

FIGURE 1: Representative two-dimensional polyacrylamide gel cyclization assays (Ulanovsky et al., 1986; Zahn & Blattner, 1988; Lyubchenko et al., 1991) for mutant M1. The ligation mixture of circles and linear molecules was separated first on 4% polyacrylamide gel (first dimension) and then on 10% polyacrylamide gel (second dimension). In the second dimension, linear molecules move on the diagonal, while circular molecules move much more slowly and are well resolved from each other and from linear species. The number of 21-bp elements (21-mers) in each oligomer species is indicated. The two gels show results for ligation in the presence (top) and the absence (bottom) of Cro. The yield of rings is significantly lower for ligation in the absence of protein.

structural changes (Mossing & Sauer, 1990; Torigoe et al., 1991), is shown. Run parameters were the following: slit width, 2.0 nm; time constant, 1 s; scan speed, 20 nm/s. DNA concentrations were calculated from absorbance spectra obtained before the CD measurements on a Gilford Response II spectrophotometer.

RESULTS AND DISCUSSION

We have used comparative cyclization determinations on a selected set of point mutations and single-base mismatches of the O_R3 site in the presence and absence of Cro protein to obtain evidence for DNA flexibility. We looked especially at alternating pyrimidine–purine sequence elements since both experimental (Steitz, 1990; Schultz et al., 1991; Donlan & Lu, 1992; McNamara et al., 1990; Bolshoy et al., 1991; Timsit et al., 1992) and theoretical (Zhurkin et al., 1979, 1991; Ulyanov & Zhurkin, 1984; Takeda et al., 1989; Zakrzewska, 1991) studies have shown these to have unusual features. We report evidence for unusual flexibility at alternating CA, CAC, and CACA sequences located in the operator regions. Parallel circular dichroism (CD) studies show that Cro binding to

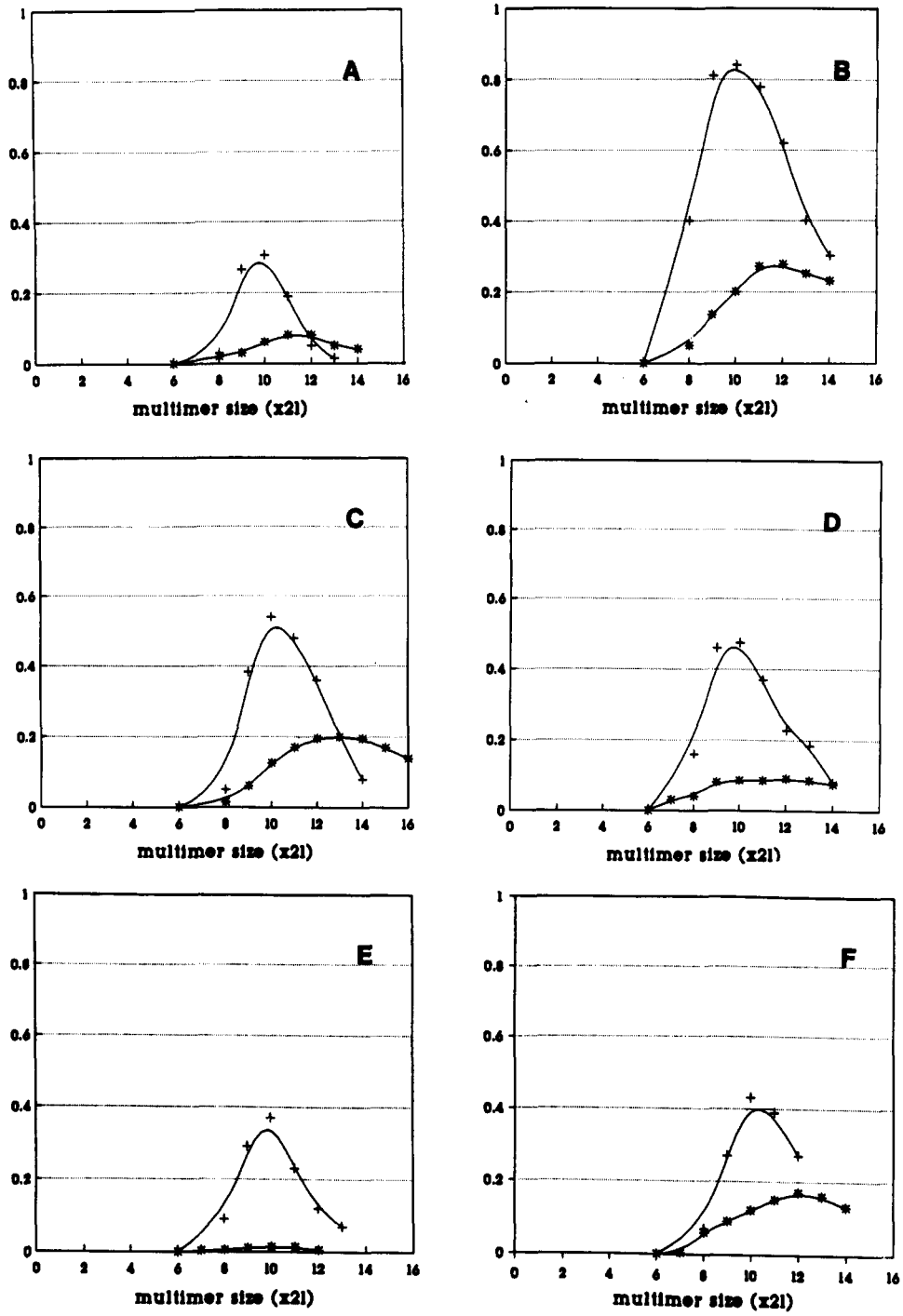

FIGURE 2: Distributions of cyclization efficiency (ordinate) as a function of fragment size (abscissa) for all mutations and mismatches obtained from the ligation of the respective 21-mers in the presence (+) and in the absence (*) of Cro protein. Cyclization efficiency determinations are described in the text. Each distribution shown is the result of at least four independent experiments. The "worst case" (i.e., large circles) experimental error in the ordinate is estimated as ±0.08 (A) Wild-type O_R3 operator sequence. (b) Mutant M1. (C) Mutant M2. (D) Mutant M3. (E) Mutant M4. (F) Mismatches COM1–3.

O_R3 induces overwinding in the DNA near the middle of the operator region and that CA elements may also facilitate this effect.

Ligation products of labeled 21-bp oligonucleotides containing 17-bp wild-type and mutated or mismatched O_R3 operator sequences as shown in Table I were analyzed by two-dimensional gel electrophoresis in the presence and absence of bound Cro protein (Lyubchenko et al., 1991). Typical gels are shown in Figure 1 (for mutant sequence M1), and cyclization data for all sequences shown in Table I are summarized in Figure 2. The maxima in the ring formation efficiency distributions shown in Figure 2 are always much greater when ligation occurs in the presence of Cro, and the distributions themselves are distinctively different. Ligation–cyclization experiments in the presence of Cro were always conducted with Cro concentrations high enough that all O_R3 sites were fully saturated with protein; small reductions in Cro concentration below those used to obtain Figure 2 had no measurable effect upon the cyclization efficiency distributions, but a factor of 2 decrease resulted in an observable shift in the maximum to larger multimer sizes (data not shown). The latter reflected a significant fraction of unbound and therefore uncurved, O_R3 oligamers in the ligation–cyclization reaction.

Certain characteristics of these results provide useful insights. The fact that bound Cro increases the cyclization

Table II: Retardation Factors for Linear 7-mer (147-bp, R_{147}) and 10-mer (210-bp, R_{210}) for the Sequences Shown in Table I

	sequence						
	WT	M1	M2	M3	COM1	COM2	COM3
R_{147}	1.15 ± 0.1	1.18 ± 0.1	1.22 ± 0.1	1.17 ± 0.1	1.15 ± 0.1	1.20 ± 0.1	1.20 ± 0.1
R_{210}	1.20 ± 0.1	1.19 ± 0.1	1.26 ± 0.1	1.20 ± 0.1	1.20 ± 0.1	1.25 ± 0.1	1.25 ± 0.1

efficiency of all sequences means that the protein bends not only the wild-type O_R3 sequence [as shown earlier in Lyubchenko et al. (1991)] but all the mutated and mismatched sequences as well. In addition, the *maxima* in the distributions for complexed oligonucleotides occur at the same fragment size irrespective of sequence. This shows that the bending angle induced by the protein is the same for all sequences. Its value (and therefore the overall geometry of DNA in the nucleoprotein complex) is determined by the protein only and not by the operator sequence. However, the operator sequence does affect the maximum cyclization efficiency, i.e., the *yield* of optimum-sized rings. This varies from 28% for wild-type O_R3 (Figure 2A) to over 80% for the M1 mutant (Figure 2B). Similar variations are evident in the cyclization efficiency distributions for the oligonucleotides ligated in the absence of Cro. In this case, the discrepancy is even greater and ranges from ~1% for the M4 mutant (Figure 2E) to 30% for the M1 mutant (Figure 2B).

To interpret these cyclization efficiency differences, it is instructive to compare the static structures, using a standard gel retardation assay for the various sequences in the absence of Cro binding. Retardation factors, R_L, defined as the ratio of *apparent* to *true* length (Lyubchenko et al., 1991; McNamara et al., 1990; Bolshoy et al., 1991), were obtained for 7-mers (147 bp) and 10-mers (210 bp) using 8% polyacrylamide gels (Lyubchenko et al., 1991; Bolshoy et al., 1991). Results are shown in Table II. The R_L values are the same within experimental uncertainty (1.2 ± 0.1). This indicates that all the free DNA sequences are very close to one another in fixed bending properties. Available compilations of sequence-dependent DNA twist angles (Kabsch et al., 1982) predict that the free DNA sequences are also closely similar in overall helical twist. Among the wild-type and mutant sequences shown in Table I, estimated differences in average base pair twist (Kabsch et al., 1982) are much less than the thermal fluctuation in this value (Zhurkin et al., 1991; Crothers et al., 1992). In addition, no temperature dependence of the R_L values was observed; this is consistent with negligible out-of-plane bending in any of these sequences (Shlyakhtenko et al., 1990). Finally, computer modeling of the wild-type and M1–M4 mutant sequences using the experimental wedge angles of Bolshoy et al. (1991) shows approximately the degree of fixed bending expected (McNamara & Harrington, 1991) from the R_L values of Table II but no significant out-of-plane bending in any of these sequences. We conclude therefore that the sequence changes investigated here, which are all *single-base substitutions*, do not significantly alter the torsional matching of ends.

The observed differences in cyclization efficiency must therefore be mainly due to flexibility differences among the 21-mers. This is also supported by the results for single base mismatched sequences (Table I, COM1–COM3). Both thermodynamic (Aboul-ela et al., 1985) and NMR (Patel et al., 1987; Patel et al., 1991, and references therein) studies have shown that mismatches destabilize the double helix and lead to structural perturbations that extend several base pairs into the duplex (Patel et al., 1991). This destabilization leads to an increase in both torsional and bending flexibility. The cyclization efficiencies of the mistmatched sequences in Figure 2F are significantly greater than the O_R3 wild-type sequence in both the presence and the absence of Cro.

A comparison of the cyclization efficiency distributions for free DNA in panels A–E of Figure 2 (curves defined by asterisks) shows that the alternating pyrimidine–purine elements CA/TG, CAC/GTG, and CACA/TGTG are potential loci of increased flexibility.

The M1 mutation shows the greatest enhancement in cyclization efficiency over wild-type O_R3 and all the mutations reported (Figure 2B). M1 differs from the wild-type sequence in a CACA/TGTG at the center of the operator. Recent crystallographic results have shown that alternating $(CA)_n(TG)_n$ elements can form an alternative structure which can propagate into adjacent sequence regions (Timsit et al., 1992). We suggest that such structural distortion leads to unusually high flexibility in the central CACA/TGTG element or possibly in the longer sequence, CACCACA, which spans the upstream binding and central operator regions and leads to the observed increase in ring closure efficiency.

The next greatest enhancement occurs with the M2 mutation (Figure 2C), which contains a single CAC/GTG element at position 4. The presence of a second CA/TG element at position 9 in the wild-type sequence leads to a large reduction in cyclization efficiency for the wild-type sequence (Figure 2A). Since the two CA/TG elements are separated by one-half helical turn in the wild-type sequence, its lower cyclization efficiency indicates that both these elements somehow compensate for each other. This suggests anisotropic flexibility (i.e., directional, as a hinge) at CA/TG and CAC/GTG. The higher efficiency of ring formation by the M2 mutant can thus be explained as anisotropic bending at the upstream CAC/GTG triplet. This concept is supported by NMR studies (Cheung et al., 1984; Donlan & Lu, 1992; Patel et al., 1987) which suggest an unusual structure for the CAC/GTG triplet, possibly involving a partial unstacking of bases in one strand (Patel et al., 1987).

Mutant M3 contains an additional GTG/CAC element at position 12. This element is a little less than one helical turn from the CAC/GTG element at position 4 and a little less than one-half turn from the CA/TG element at position 9. If the anisotropy of bending is similar among these three elements, anisotropic flexing at position 13 might partially compensate for other bending loci but could not fully counteract such bending at position 4 or fully reinforce it at position 9. Consistent with this picture, the M3 mutation has a cyclization efficiency sharply reduced compared to mutants M1 and M2 but roughly comparable to the wild-type O_R3 sequence.

Finally, if increased flexibility is associated with elements comprising alternating CA elements, the absence of all such elements should lead to low cyclization efficiency. This is clearly demonstrated by the results for mutant M4 (Figure 2E), for which the maximum efficiency is below 1%. A comparison of the free DNA cyclization efficiencies in panels B–E of Figure 2 suggests that flexibilities are in the order CACA/TGTG > CAC/GTG, CA/TG. These data cannot determine the direction of bending at the flexible elements. However, Monte-Carlo calculations (Zhurkin et al., 1979, 1991; Ulyanov & Zhurkin, 1984) indicate a preference for bending into the major groove.

Insights into the role of the operator sequence on Cro protein binding can be obtained from comparisons of cyclization

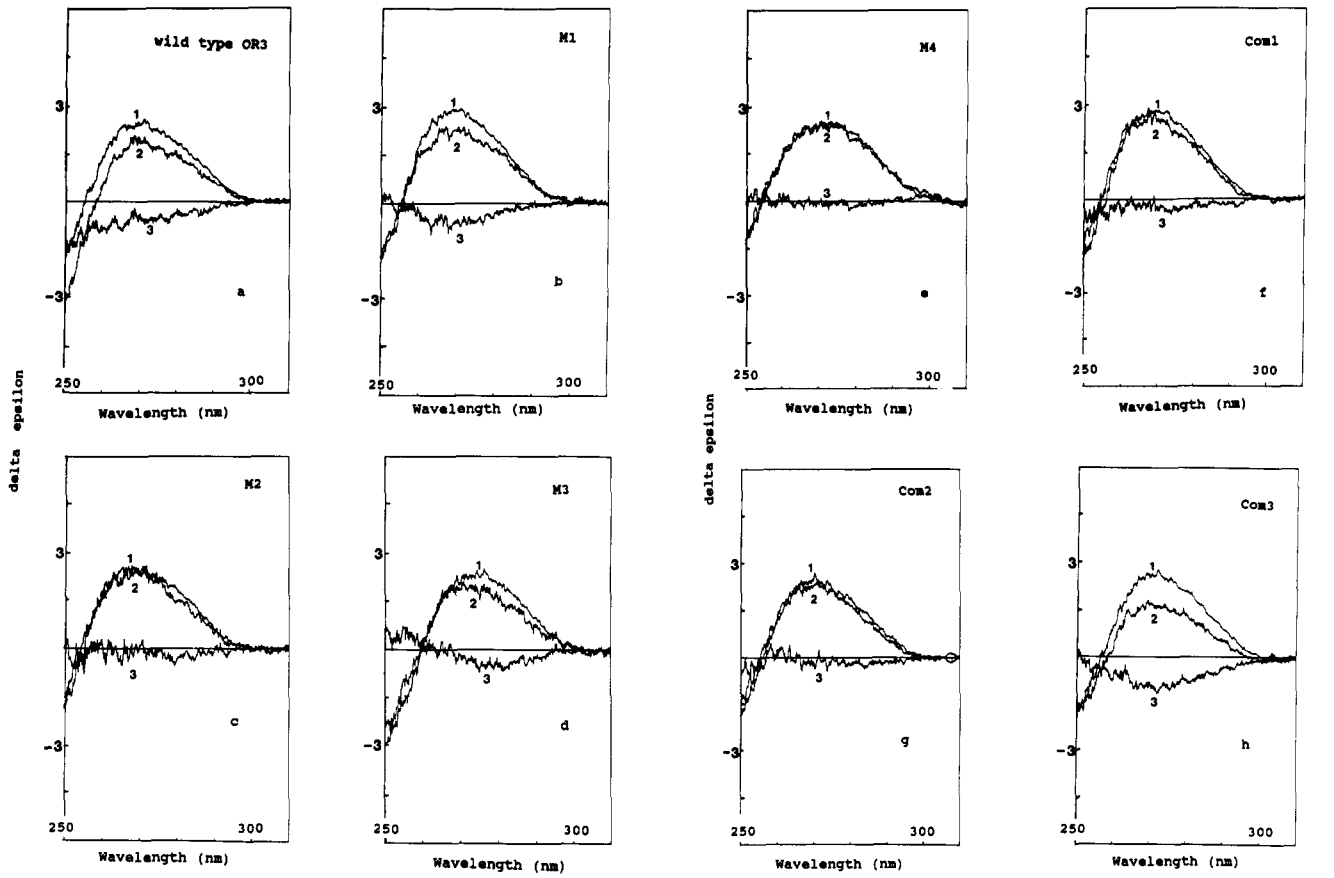

FIGURE 3: Circular dichroism (CD) spectra. Each panel shows CD spectra of 21-bp oligonucleotides alone (1) and complexes of the oligonucleotides with the Cro protein (2) and different spectra (3) between (2) and (1). Each spectrum shown is the average of at least three independent spectra obtained under the same run conditions. DNA concentrations were calculated from absorbance spectra obtained before the CD measurements on a Gilford Response II spectrophotometer. (a) Wild-type sequence. (b) Mutant M1. (c) Mutant M2. (d) Mutant M3. (e) Mutant M4. (f) Mismatch COM1. (g) Mismatch COM2. (h) Mismatch COM3. Spectra were obtained by scanning between 210 and 310 nm, but only the longer wavelength part of each spectrum, which reflects DNA structural changes (Mossing & Sauer, 1990; Torigoe et al., 1991) is shown.

efficiency distributions in the presence of Cro. These are shown in Figure 2 (curves defined by pulses). Mutations M1 and M2 occur in the central region of the O_R3 operator sequence, which is not involved in specific interactions with the protein (Metzler & Lu, 1989; Takeda et al., 1989; Brennan et al., 1990; Hochshild & Ptashne, 1986). The cyclization efficiency of both mutations is considerably enhanced over wild-type O_R3. The effect is especially pronounced for mutation M1, which shows a 3-fold increase, indicating that the structure associated with CACA/TGTG in the central region, and possibly its propagation into the CAC/GTG element at position 4, is responsible. Since the position of the distribution maximum and hence the bending angle are not changed, this is most likely a result of increased flexibility. A recent crystallographic study has shown that the CACA motif in the crystal at low temperatures can adopt an unusual structure with frame-shifted A·G and T·C interstrand hydrogen bonds and considerable base tilt (Timsit et al., 1992). Transformation of the central CACA element into this type of structure would almost certainly lead to increased flexibility at this point which might facilitate DNA bending by the Cro protein. On the other hand, mutation M2 has lost the CA at position 9 in the central region. Its cyclization efficiency is appreciably higher than that of the wild-type O_R3, although it is less than that of the M1 mutant. Comparison of cyclization efficiencies for free M2 and wild-type O_R3 sequences suggests increased anisotropic flexibility at the CA/TG dinucleotide. In addition, a comparison of cyclization data between free DNA and complex for the same sequences suggests that bending at the CA/TG element in the central region is opposite to that induced by Cro. Thus, the CA/TG element in the central region appears to be the sequence feature primarily responsible for the decreased cyclization efficiency observed in the wild-type O_R3 operator. It may function to decrease Cro-induced DNA bending since a single CA/TG element in the central region leads to the lowest cyclization efficiency, but this bending is dramatically increased by its removal or its amplification into CACA/TGTG.

The GTG/CAC element centered at position 12 in the M3 mutant is located in the downstream specific binding region (Metzler & Lu, 1989; Takeda et al., 1989). In the presence of Cro, the cyclization efficiency of M3 increases almost 2-fold over that of wild-type O_R3, while the efficiencies for free DNA are almost identical. The observed enhancement must therefore be due entirely to increased flexibility effects which derive from interaction with the protein. The M4 mutant differs from M2 only in the removal of the CAC/GTG at position 4, but it nevertheless shows reduced cyclization in the presence of Cro, comparable roughly to the wild-type O_R3 sequence. This mutant therefore seems to show bending effects due to the Cro protein with only minor contributions from inherent DNA flexibility. These comparisons show that CAC/GTG elements in the specific binding regions have significant effects on operator flexibility as manifested in enhanced cyclization efficiency. A still larger effect is observed with mutation M1, which contains CACCACA spanning the upstream binding and central operator regions.

The gel experiments allow us to investigate Cro-induced DNA bending and disclose the effects of increased flexibility in the operator region DNA on Cro binding. Parallel circular

dichroism (CD) studies demonstrate that overwinding occurs in the central nondirect binding operator region as well. CD results are shown in Figure 3. Scans were made from 210 to 310 nm, but for clarity of presentation the spectra shown are restricted to the longer wavelength part of the spectrum which reflects changes in DNA structure only (Mossing & Sauer, 1990; Torigoe et al., 1991). Intensity of the positive CD band at 275 nm depends upon the helix winding angle (Tunis-Schneider & Maestre, 1970; Johnson et al., 1981; Torigoe et al., 1991) such that the lower the intensity, the larger the winding angle. The spectra for the pure oligonucleotides (curve 1 in all panels) are characteristic of B-form DNA (Johnson et al., 1981; Torigoe et al., 1991), but the spectrum of the M1 oligonucleotide shows underwinding in comparison with the wild type and the other mutations. This could be explained by a small contribution from an A-form type of structure, shown recently for CACA-containing oligonucleotides (Timsit et al., 1992; Jain et al., 1989). Formation of A·G and T·C hydrogen bonds in the CACA elements as suggested by Timsit et al. (1992) would lead to distortion and inclination of AT and GC base pairs in a manner similar to A-form DNA (Olsen et al., personal communication). Complexing with Cro decreases the amplitude at 275 nm (curve 2 on Figure 3a), and the difference spectra are negative in this region. This is characteristic of DNA overwinding in the nucleoprotein complex.

The pronounced difference spectra in panels a, b, and d of Figure 3 indicate DNA overwinding by the Cro protein bound to the wild-type O_R3 and to the M1 and M3 mutations. No comparable difference spectra are observed for mutations M2 (Figure 3c) or M4 (Figure 3e), and Cro does not change the DNA winding in either of these sequences. The sequences showing difference spectra have a CA element at position 9, and M1 has an additional CA at position 8. All these are in the central operator region. The M2 and M4 mutations both entail a G-for-C substitution in the central operator region that eliminates the central CA element at position 9. Thus, the *central region* of the operator DNA appears to be sensitive to overwinding by Cro binding. This conclusion is further supported by the spectra of mismatches: no overwinding by Cro occurs for COM2 (Figure 3g), where a G/G mismatch is located in the central region, or for COM1 (Figure 3f), where an A/C mismatch is located at position 8. These mismatches both occur in the central region. The absence of overwinding in these sequences can therefore be understood if the overwinding induced by the Cro protein is relieved by flexibility due to the mismatch. On the other hand, considerable overwinding is evident in mismatch COM3 (Figure 3h), which has a T/G mismatch at position 13 in the downstream protein binding region. Thus, the CD results clearly show that Cro binding overwinds DNA in the central operator region between positions 5 and 13. This correlates with crystallographic data (Brennan et al., 1990) and may help explain NMR results (Kirpichnikov et al., 1964; Metzler & Lu, 1989; Lee et al., 1987) on the Cro–O_R3 complex. The CD data also suggest that the CA sequence element may facilitate this overwinding, although the possibility that a purine-rich central region, as in the M2 and M4 mutants, is characterized by a more rigid structure (Zhurkin et al., 1979, 1991; Ulyanov & Zhurkin, 1984; Zakrzewska, 1990), which resists protein-induced DNA twisting, cannot be excluded.

In this work, we have shown that the alternating pyrimidine–purine elements CA/TG, CAC/GTG, and CACA/TGTG are sites of exceptional flexibility in DNA sequences derived from the λ phase O_R3 site and in the complexes of these sequences with the Cro protein. The CAC/GTG element in the specific binding region of Cro with the O_R3 operator evidently facilitates protein-induced DNA bending. This is consistent with the observation that two CAC/GTG elements at the O_L1 and O_R2 sites improve interaction with both Cro and CI repressor proteins (Kabsch et al., 1982; Sarai & Takeda, 1989). In addition, the CA element in the central region may influence DNA overwinding.

The triplet CAC/GTG is a ubiquitous sequence feature in gene regulatory regions (Donlan & Lu, 1992; McNamara et al., 1990; Barber & Zhurkin, 1990) and has been proposed as a potential site of alternative DNA structure (Cheung et al., 1984; Donlan & Lu, 1992; McNamara et al., 1990; Bolshoy et al., 1991; Timsit et al., 1992). The CAC/GTG element appears in the mouse β-globin promoter (Cowie & Myers, 1988), and sequence motifs CACGTG/GTGCAC and CACAGTG/GTGTCAC appear as consensus regulatory elements in immunoglobulin and T-cell receptor genes (Hesse & Gellert, 1989), in recognition elements of *myc* oncogenes (Kerkhoff et al., 1991), and in plant G-box regions (Gilmartin et al., 1990). Earlier theoretical work has suggested enhanced flexibility in alternating pyrimidine–purine DNA sequences (Zhurkin et al., 1979; Ulyanov & Zhurkin, 1984). The CA/TG element in the *CAP* operator site has been shown to kink through a large angle in the specific binding complex with the protein (Steitz, 1990; Schultz et al., 1991). DNA bending has also been shown to accompany the noncooperative binding (Hochshild & Ptashne, 1986) of the Cro protein of λ phage to the 17-bp O_R3 operator site (Lyubchenko et al., 1991; Brennan et al., 1990). The results of this study suggest a physical basis for these various observations.

ACKNOWLEDGMENT

We thank Y. Takeda for a generous gift of Cro protein and D. Shlyakhtenko (Department of Mathematics, University of Nevada—Reno) and D. Rekesh (Department of Physics, Arizona State University) for specialized computer software used respectively for CD data processing and molecular graphics. The authors are also indebted to M. Gellert, B. W. Matthews, M. C. Mossing, E. N. Trifonov, P. H. von Hippel, and V. B. Zhurkin for helpful and stimulating discussions.

REFERENCES

Aboul-ela, F., Koh, D., Tinoko, I., & Martin, F. (1985) *Nucleic Acids Res. 13*, 4811–4823.

Barber, A. M., & Zhurkin, V. B. (1990) *J. Biomol. Struct. Dyn. 8*, 213–232.

Bhattacharyya, A., & Lilley, D. M. J. (1989) *Nucleic Acids Res. 17*, 6821–6840.

Bolshoy, A., McNamara, P. T., Harrington, R. E., & Trifonov, E. N. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2312–2316.

Brennan, R. G., Roderick, S. L., Takeda, Y., & Matthews, B. W. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8165–8169.

Cheung, S., Arndt, K., & Lu, P. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3665–3669.

Cowie, A., & Myers, R. M. (1988) *Mol. Cell. Biol. 8*, 3122–3128.

Crothers, D. M., Drak, J., Kahn, J. D., & Levene, S. D. (1992) *Methods Enzymol. 212*, 3–29.

Donlan, M. E., & Lu, P. (1992) *Nucleic Acids Res. 20*, 525–532.

Drew, H. R., & Travers, A. A. (1985) *Nucleic Acids Res. 13*, 4445–4467.

Gilmartin, P. M., Sarokin, L., Memelink, J., & Chua, N.-H. (1990) *Plant Cell 2*, 369–378.

Hagerman, P. J., & Ramadevi, V. A. (1990) *J. Mol. Biol. 212*, 351–362.

Harrison, S. C., & Aggarvwal, A. K. (1990) *Annu. Rev. Biochem. 59*, 933–969.

Hesse, J. E., & Gellert, M. (1989) *Genes Dev. 3*, 1053–1061.

Hochshild, A., & Ptashne, M. (1986) *Cell 44*, 925–933.
Horowitz, D. S., & Wang, J. C. (1984) *J. Mol. Biol. 173*, 75–91.
Hsien, C.-H., & Griffith, J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4833–4837.
Husain, I., Griffith, J., & Sancar, A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2558–2562.
Jain, S., Zon, G., & Sundaralingam, M. (1989) *Biochemistry 28*, 2360–2364.
Johnson, B. B., Dahl, K. S., Tinoco, I., Jr., & Ivanov, V. I. (1981) *Biochemistry 20*, 73–78.
Kabsch, W., Sander, S., & Trifonov, E. N. (1982) *Nucleic Acids Res. 10*, 1097–1104.
Kerkhoff, E., Blister, K., & Klempnauer, K.-H. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4323–4327.
Kirpichnikov, M. P., Hahn, K. D., Buck, F., Ruterjans, H., Chernov, B. K., Kurochkin, A. V., Skryabin, K. G., & Bayev, A. A. (1964) *Nucleic Acids Res. 12*, 3551–3561.
Lee, S. J., Shirakava, M., Akustsu, H., Kyogoku, Y., Shiraishi, M., Kitano, K., Shin, M., Ohtsuka, E., & Ikehara, M. (1987) *EMBO J. 6*, 1129–1135.
Leroy, J. L., Kochoyan, M., Huynh-Dinh, T., & Gueron, M. (1988a) *J. Mol. Biol. 200*, 223–238.
Leroy, J. L., Charretier, E., Kochoyan, M., & Gueron, M. (1988b) *Biochemistry 27*, 8894–8898.
Levene, S. D., & Crothers, D. M. (1986a) *J. Mol. Biol. 189*, 61–72.
Levene, S. D., & Crothers, D. M. (1986b). *J. Mol. Biol. 189*, 73–83.
Lyubchenko, Y. L., Shlyakhtenko, L. S., Chernov, B. K., & Harrington, R. E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 5331–5334.
McNamara, P. T., & Harrington, R. E. (1991) *J. Biol. Chem. 286*, 12548–12554.
McNamara, P. T., Bolshoy, A., Trifonov, E. N., & Harrington, R. E. (1990) *J. Biomol. Struct. Dyn. 8*, 529–538.
Metzler, W. J., & Lu, P. (1989) *J. Mol. Biol. 205*, 149–164.
Mossing, M. C., & Sauer, R. T. (1990) *Science 250*, 1712–1715.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984) *Fed. Proc. 43*, 2663–2670.
Patel, D. J., Shapiro, L., & Hare, D. (1987) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds. Springer-Verlag, New York, pp 115–161.
Sarai, A., & Takeda, Y. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6513–6517.
Schellman, J. A. (1974) *Biopolymers 13*, 216–226.
Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001–1007.
Shlyakhtenko, L. S., Lyubchenko, Y. L., Chernov, B. K., & Zhurkin, V. B. (1990) *Mol. Biol. (Engl. Transl.) 24*, 66–81.
Shore, D., & Baldwin, R. L. (1983a) *J. Mol. Biol. 170*, 957–981.
Shore, D., & Baldwin, R. L. (1983b) *J. Mol. Biol. 170*, 983–1007.
Steitz, T. A. (1990) *Q. Rev. Biophys. 23*, 205–280.
Takeda, Y., Sakai, A., & Rivera, V. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 439–443.
Taylor, W. H., & Hagerman, P. J. (1990) *J. Mol. Biol. 212*, 363–376.
Timsit, Y., Vilbois, E., & Moras, D. (1992) *Nature 354*, 167–170.
Torigoe, C., Kidokoro, S., Takimoto, M., Kyogoku, Y., & Wada, A. (1991) *J. Mol. Biol. 219*, 733–746.
Travers, A. A. (1989) *Annu. Rev. Biochem. 58*, 427–452.
Tunis-Schneider, M.-J. B., & Maestre, M. F. (1970) *J. Mol. Biol. 52*, 521–541.
Ulanovsky, L. E., Bodner, M., Trifonov, E. N., & Choder, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 862–866.
Ulyanov, N. B., & Zhurkin, V. B. (1984) *J. Biomol. Struct. Dyn. 2*, 361–385.
Zahn, K., & Blattner, F. R. (1988) *Science 236*, 416–422.
Zakrzewska, K. (1992) *J. Biomol. Struct. Dyn. 9*, 681–693.
Zhurkin, V. B. (1985) *J. Biomol. Struct. Dyn. 2*, 785–804.
Zhurkin, V. B., Lysov, Y. P., & Ivanov, V. I. (1979) *Nucleic Acids Res. 6*, 1081–1096.
Zhurkin, V. B., Ulyanov, N. B., Gorin, A. A., & Jernigan, R. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7046–7050.