

## Design and Synthesis of Sequence-Specific DNA-Binding Peptides

S.L. Grokhovsky, A.N. Surovaya, R.V. Brussov,  
B.K. Chernov, N.Yu. Sidorova and G.V. Gursky

W.A. Engelhardt Institute of Molecular Biology  
Academy of Sciences of the USSR  
Vavilov str. 32  
Moscow 117984, USSR

### Abstract

Design, synthesis and DNA binding activities of two peptides containing 32 and 102 residues are reported. A nonlinear 102-residue peptide contains four modified  $\alpha$  helix-turn- $\alpha$  helix motifs of 434 *cro* protein. These four units are linked covalently to a carboxyterminal crosslinker containing four arms each ending with an aliphatic amino group. From CD studies we have found that in aqueous buffer in the presence of 20% trifluoroethanol the peptide residues assume  $\alpha$ -helical,  $\beta$ -sheet and random-coiled conformations with the  $\alpha$ -helical content of about 16% at room temperature. Upon complex formation between peptide and DNA, a change in the peptide conformation takes place which is consistent with an  $\alpha$ - $\beta$  transition in the DNA binding  $\alpha$  helix-turn- $\alpha$  helix units of the peptide. Similar conformation changes are observed upon complex formation with the synthetic operator of a linear peptide containing residues 7-37 of 434 *cro* repressor. Evidently, in the complex, residues present in helices  $\alpha_2$  and  $\alpha_3$  of the two helix motif form a  $\beta$ -hairpin which is inserted in the minor DNA groove. The last inference is supported by our observations that the two peptides can displace the minor groove-binding antibiotic distamycin A from poly(dA)·poly(dT) and synthetic operator DNA. As revealed from DNase digestion studies, the nonlinear peptide binds more strongly to a pseudooperator  $O_{p1}$ , located in the *cro* gene, than to the operator  $O_{R3}$ . A difference in the specificity shown by the non-linear peptide and wild-type *cro* could be attributed to a flexibility of the linker chains between the DNA-binding domains in the peptide molecule as well as to a replacement of Thr-Ala in the peptide  $\alpha_2$ -helices. Removal of two residues from the N-terminus of helix  $\alpha_2$  in each of the four DNA-binding domains of the peptide leads to a loss of binding specificity.

### Introduction

There is much interest in design and synthesis of sequence-specific DNA-binding peptides. These peptides could be used to affect selectively the activity of certain genes in bacterial and eucaryotic cells and might have important implications as antiviral and antitumor agents.

At present it is well-known that relatively short sequences in gene-regulating proteins contain all (or almost all) of the information needed for the recognition of their



A synthetic 52-  
 inase and con-  
 ectively to *Hin*  
 n be converted  
 ent of an iron  
 nus (3,4).

idues can bind  
 peptides offers  
 tides. It should  
 osomes in bac-

ng activity of a  
 our modified  $\alpha$   
 valently linked

to the C-terminal crosslinker which contains four arms each ending with an aliphatic amino group. The lengths of the linker chains between the DNA binding domains have been chosen to extend over six DNA base pairs, a distance known to separate the DNA binding domains in the two symmetry-related subunits of 434 *cro* (or 434 repressor) complexed with synthetic operators (10,11). From molecular model building studies it can be easily seen that a tetrahedral disposition of the DNA binding domains becomes possible in the peptide I molecule, in a close similarity with that observed in a protein tetramer.

Physico-chemical studies have shown that DNA binding domains of peptide I to exist in a disordered conformation in aqueous solution and to assume  $\alpha$ -helical and  $\beta$ -sheet conformations in the presence of 20% (v/v) trifluoroethanol (TFE). At room temperature in the presence of 20% TFE the content of  $\alpha$ -,  $\beta$ - and disordered conformation has been estimated to be 0.16, 0.21 and 0.63, respectively (8). Peptide I forms very tight complexes with DNA which are stable in the presence of 1 M  $\text{NH}_4\text{F}$ . However, we have failed to observe selective binding of the peptide I to three non-identical operator sites recognized by 434 repressor and *cro* proteins in the rightward operator ( $\text{O}_R$ ) (9). These results deserve some comment. The DNA-binding domains of peptide I contain practically all amino acid residues present in the recognition helix  $\alpha_3$  of the 434 *cro*. Only one replacement (Glu37-Thr) has been introduced. However, glutamate 37 has been shown to contact residues in helices 1 and 3 of 434 *cro* and not to participate in DNA interactions (10). Substitutions at this position of recognition helix are unlikely to affect the binding specificity of 434 repressor and *cro* proteins (10-13).

The fact that peptide I does not form specific complexes with the operator sites  $\text{O}_R1$ ,  $\text{O}_R2$  and  $\text{O}_R3$  can be explained by two reasons:

1. Chemical structure of the nonlinear peptide does not allow for the two modified helix-turn-helix motifs of 434 *cro* to bind simultaneously to the two halves of the symmetrical operator, thereby decreasing the strength of the peptide binding to the operator DNA.

2. Specific binding of the 434 *cro* to DNA is mediated not only by residues located in helix 3, but also some residues which lie outside the recognition helix and are missing in the peptide I. In the present communication, we report data on synthesis and DNA interaction of a peptide II, which differs from the peptide I by the presence of a dipeptide Thr Gln at the N-terminus of each of the four arms (Figure 1). This dipeptide is also present at the N-terminus of helix 2 of 434 *cro*. In the crystalline 434 *cro* repressor-operator complex, the side chain and peptide NH group of glutamine 19 are implicated in interaction with DNA phosphate groups (10). Experiments described below show that peptide II binds specifically to the operator sites  $\text{O}_R1$ ,  $\text{O}_R2$  and  $\text{O}_R3$ .

Our previous CD studies have shown that binding of peptide I to DNA and synthetic polynucleotides is accompanied by considerable conformation changes in the peptide molecule that correspond to  $\alpha$ - $\beta$  transitions in the modified  $\alpha$  helix-turn- $\alpha$

III

434 Cro

of phage 434 *cro*  
 reviations for the  
 In: R, Arg, S, Ser;

helix motifs of 434 *cro* (8). Evidence presented here show that similar structural changes take place upon binding of peptide II to calf thymus DNA and synthetic operator DNA fragment.

In view of these findings, we decided to synthesize a linear peptide III (Figure 1) and to include it in further studies. The peptide contains residues 7-37 of 434 *cro* preceded by a glycine residue and fluorescent chromophore Dns (Dns is a residue of 5-dimethylamino naphthalene-1-sulfonic acid). Only three amino acid replacements (Arg12- Cys, Glu21- Ala and Glu37- Val) have been introduced. As revealed from the crystalline structure of 434 *cro*-operator complex, these three residues as well as alanine 38 in helix 3 are not implicated in specific interactions with DNA base pairs, but might be important for stabilization of the overall structure of the DNA binding domain (10). Glutamate 37 is hydrogen bonded to both arginine 12 and glutamine 19. This hydrogen bond network is thought to be responsible for keeping the helices 1, 2 and 3 of 434 *cro* in the correct relative orientation (10). We have found that peptide III exists partially in the  $\alpha$ -helical structure in the presence of 20% TFE and that it undergoes conformation changes (presumably,  $\alpha$ - $\beta$  transition) upon binding to a synthetic operator, in a close similarity with the behavior of peptide II. To estimate the affinities of the peptides II and III to synthetic operator DNA fragments competition-type experiments have been carried out using antibiotic distamycin A as a DNA-binding ligand with known binding characteristics (14).

### *Materials and Methods*

#### *Synthesis of Peptides I, II and III*

The peptide I was prepared by step-wise manual solid phase synthesis in accordance with the earlier published protocols (8,15). Initially a carboxyterminal crosslinker was synthesized. It contained three lysine residues of which one was joined through its carboxyl to phenylacetamidomethyl (PAM) resin substituted with valine. The crosslinker had four arms attached to  $\alpha$ - and  $\epsilon$ -aminogroups of other two lysine residues. In the peptide I, each arm of the crosslinker was coupled with the peptide containing slightly modified  $\alpha$ -helix-turn- $\alpha$  helix motif of phage 434 *cro* repressor (Figure 1). Each cycle of peptide solid-phase synthesis was monitored by quantitative fluoroscamine analysis (16) and couplings were repeated, until the yield was maximal. After deprotection with anhydrous hydrogen fluoride (17) the peptide was purified by gel chromatography. The data on the amino acid composition and sequence of peptide I were published elsewhere (8). Peptide II was prepared by joining glutamine and threonine residues to the protected resin-bound peptide I. Removal of the peptide from the resin and its purification were carried out following the procedures reported previously for peptide I (8). Amino acid composition of peptide II was as follows:

Aca 9.5 (10), Ala 17.9 (16), Glu 4.1 (4), Gly 15.9 (16), Ile 7.9 (8),  
Leu 8.5 (8), Lys 18.3 (19), Ser 3 (4), Thr 10.3 (12), Val 5.2 (5).

Here Aca is a residue of  $\epsilon$ -aminocaproic acid. Bracketed is the number of residues of a given type in the peptide molecule according to the chemical formula. The

milar structural  
A and synthetic

II (Figure 1) and  
7 of 434 *cro* pre-  
is a residue of 5-  
id replacements  
s revealed from  
sidues as well as  
DNA base pairs,  
e DNA binding  
2 and glutamine  
eping the helices  
und that peptide  
TFE and that it  
on binding to a  
l. To estimate the  
nts competition-  
in A as a DNA-

is in accordance  
inal crosslinker  
s joined through  
with valine. The  
other two lysine  
with the peptide  
34 *cro* repressor  
d by quantitative  
ld was maximal.  
ide was purified  
and sequence of  
ining glutamine  
val of the peptide  
cedures reported  
as as follows:

7.9 (8),  
(5).

number of residues  
al formula. The

molecular masses of neutral forms of peptide I and peptide II, containing 94 and 102 amino acid residues, are equal to 9951 and 10868 daltons, respectively. Lyophilized samples of peptide I and II, predried in the presence of phosphorus anhydride and in vacuo, were used for preparations of peptide solutions. Concentration of the peptide was determined by weighting the dried peptide samples and dissolving them in known volumes of buffer. The content of counterions in the peptide sample was taken to be equal to the number of positively charged lysine residues. Concentration of peptide II was also determined by measuring the amount of amino acid residue of a given type in a certain volume of solution after acid hydrolysis of the peptide. Peptide III was prepared by step-wise manual solid-phase method with modifications described previously (8,9). The following protective groups for amino acid side chains were used: Tos (Arg), 2ClZ (Lys), 4MeBzl (Cys) and OBzl (Ser, Thr). Pentafluorophenol esters of Gly, Lys, Cys, Ala, Met, Ser, Thr and Val residues were prepared prior to condensation and purified by recrystallization. Symmetrical anhydrides of Ile, Ser and Leu residues and N-hydroxybenzotriazole esters of Arg, Gln and Dns Gly were prepared just before condensation. After completion of peptide synthesis and deprotection with anhydrous hydrogen fluoride (17) the peptide was purified by gel chromatography (Toyopearl HW-40) in 1M CH<sub>3</sub>COOH and by reverse phase preparative HPLC (Zorbax ODS, 10 μ column (21.2 × 250 mm)) in 80% acetonitrile-water gradient in the presence of 0.1% trifluoroacetic acid. Amino acid composition of the peptide was as follows:

Ala 4.4 (4), Arg 1.9 (2), Cys (not determined), Gln 4.8 (4), Gly 1.3 (1),  
Ile 2.8 (3), Leu 4.1 (4), Met 0.8 (1), Ser 1.1 (1), Thr 3.3 (3), Val 1.7 (2).

Bracketed is the content of a specified amino acid residue expected from the chemical formula of the peptide. Concentration of the peptide was determined spectrophotometrically using a value of 4300 for the molar absorbance at 330 nm.

#### *DNA and Synthetic Polydeoxyribonucleotides*

Poly(dA) · poly(dT) and poly(dG) · poly(dC) (PL Biochemicals), poly[d(AC)] · poly[d(GT)] (Boehringer Mannheim) and calf thymus DNA (Sigma) were used as supplied, without further purification. Poly(dG) · poly(dC) was dissolved in 0.1 N NaOH and then dialyzed against 1mM sodium cacodylate buffer (pH 7.0) in the presence of 1mM EDTA. Other polydeoxyribonucleotides and DNA were dissolved in 1mM sodium cacodylate buffer and dialyzed for 24 h against the same buffer in the presence of 1mM EDTA. Prior to experiment, solutions of all the nucleic acids were dialyzed twice for 48 h against 1mM sodium cacodylate buffer (pH 7.0). The UV spectra were obtained with a Cary 2200 spectrophotometer. The CD spectra were recorded with a Jobin-Ivon Mark III dichrograph using 0.05 cm, 0.10 cm and 0.50 cm pathlength cells.

#### *Isolation of DNA Fragments Containing Phage 434 Operator Sites O<sub>R1</sub>, O<sub>R2</sub> and O<sub>R3</sub>*

The 434 bacteriophage was grown using the *Escherichia coli* strain LE392. The smallest of the fragments formed after cleavage of λ<sub>imm</sub> 434 DNA by restriction

endonuclease EcoRI (EcoRI-G, 1283 bp) was cloned in a plasmid pUC9 at the cut site of EcoRI restrictase. The bacteria were transformed by means of the HB101-RR1 strain following the procedure reported elsewhere (18,19). The plasmid was isolated by lysis of the bacterial cells at alkaline pH, purified by ultracentrifugation in a cesium chloride/ethidium bromide density gradient, extracted with n-butanol to remove ethidium bromide, dialyzed against low-salt buffer to remove cesium chloride, concentrated by ethanol precipitation and redissolved in 10 mM Tris-HCl buffer pH 7.5, 0.1 mM EDTA. The purified plasmid was then digested with EcoRI endonuclease to leave the EcoRI-G fragment and linearized plasmid. From this mixture the DNA fragment was separated by electrophoresis on 5% polyacrylamide gel and digested with restriction endonucleases Sau3A and HpaII. The two DNA fragments, A and B, containing 306 and 217 bp, respectively, were isolated by electrophoresis on 5% polyacrylamide gel. The fragment C (410 bp) was isolated from a pUC9 plasmid that was digested with restriction endonucleases BspRI and HinfI. The bottom and top strands of fragments A and B were 3' end-labeled by end-filling with [ $\alpha$ - $^{32}$ P] dATP and [ $\alpha$ - $^{32}$ P] dCTP and DNA polymerase Klenow fragment (Boehringer Mannheim) at the EcoRI and HpaII cut sites, respectively. Incubation of DNA fragment C with DNA polymerase I Klenow fragment and [ $\alpha$ - $^{32}$ P] dGTP led to selective radioactive labeling of the 3' end of the fragment at the HinfI site.

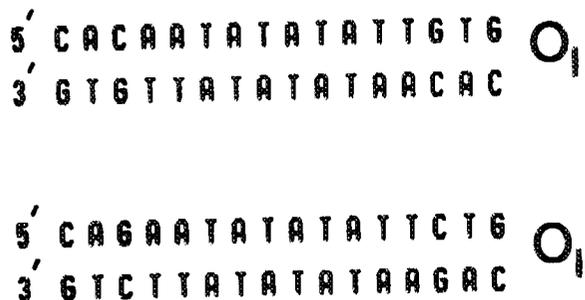
*Preparation of DNA Complexes with Peptides and Partial Digestion of Naked DNA and Peptide-DNA Complexes by Acidic DNase and Pancreatic DNase I*

For experiments using acidic DNase isolated from wolleye pollock liver (20) samples (5  $\mu$ l) of the labeled DNA fragment ( $10^4$  counts per minute) and unlabeled pUC9 DNA at concentration of 15-70  $\mu$ M (base pairs) were mixed with 15  $\mu$ l of peptide solution (0.4-40.0  $\mu$ M) in 10 mM Tris-HCl buffer (pH 7.5) containing 30% (v/v) TFE. The mixture containing 20% TFE was kept for 20 min at 20°C, after which time 2  $\mu$ l of 0.3 M sodium acetate (pH 5.5) and 2-4  $\mu$ l of acidic DNase at concentration of 6-25  $\mu$ g/ml were added. Digestion was proceeded for 60 seconds at 20°C. The reaction was stopped by adding 100  $\mu$ l of solution containing 50 mM Tris-HCl buffer (pH 7.5), 0.15 M NaCl, 10 mM EDTA and 10  $\mu$ g/ml of tRNA. The DNA was extracted with phenol, precipitated with ethanol, washed in 70% ethanol, dried and analyzed on 6% polyacrylamide denaturing gel 40 cm long with the gradient thickness of 0.15-0.3 mm (21).

For experiments using DNase I as a DNA-cleaving agent, 5  $\mu$ l of the labeled restriction fragment ( $10^4$  counts per minute) and unlabeled pUC9 DNA at concentration of  $7 \cdot 10^{-5}$  M (base pairs) were incubated with 5  $\mu$ l of peptide solution at concentration of 5-20  $\mu$ M in a buffer containing 10 mM Tris-HCl (pH 7.5), 100 mM NaCl and 15% TFE. The mixture was kept for 20 min at 20°C and then incubated for 4 min with 2  $\mu$ l of DNase I in the same buffer containing 5 mM MgCl<sub>2</sub> (final enzyme concentration of 16  $\mu$ g/ml). Reactions were terminated by adding the stop solution identical to that used for termination of DNA cleavage by acidic DNAase.

*Synthesis of Operator DNA Fragments O<sub>I</sub> and O<sub>II</sub>*

Synthetic operator DNA fragments O<sub>I</sub> and O<sub>II</sub> (Figure 2) contain selfcomplementary

Figure 2: Chemical structures of synthetic operators  $O_I$  and  $O_{II}$ .

oligonucleotides which were prepared by a manual modification of solid phase amidophosphate method (22). After deprotection and removal of all the protective groups the two oligonucleotides were isolated and purified by anion exchange and reverse phase high pressure liquid chromatography. Nucleotide sequences of the synthetic oligonucleotides were confirmed by method of Maxam and Gilbert (21).

### Results

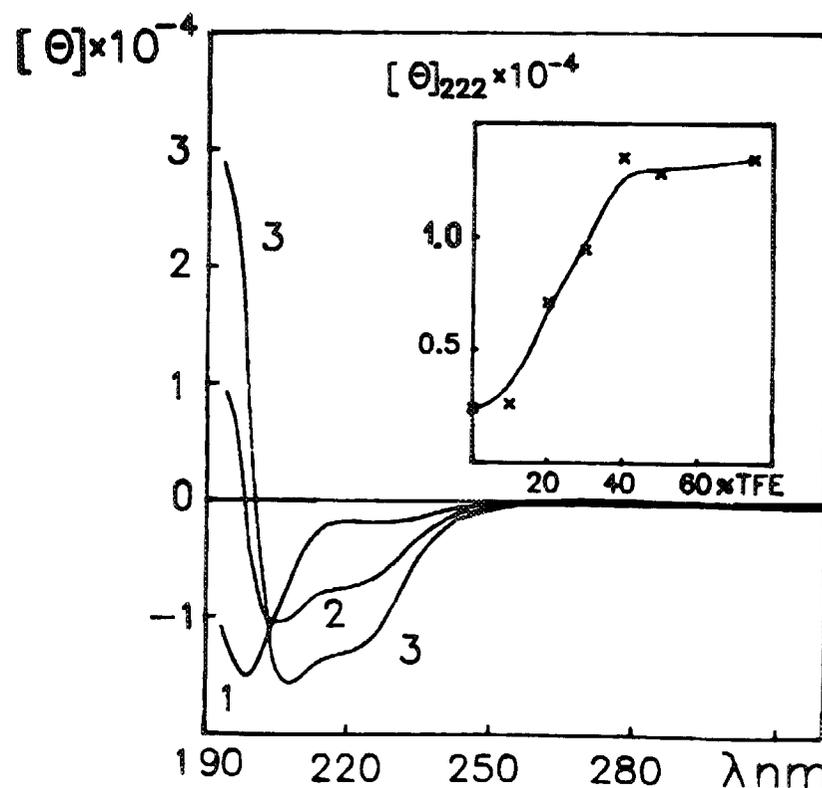
#### CD Spectra of Peptide II in Water-Trifluoroethanol Mixtures

Figure 3 shows CD spectra of peptide II in aqueous buffer in the presence of various concentrations of TFE. The CD spectrum recorded for peptide II in aqueous buffer without TFE shows a negative band near 200 nm with a shoulder near 220 nm. This indicates that a significant part of the peptide residues are in a disordered conformation. In the presence of 20% (v/v) TFE the shape of CD spectrum is drastically changed. Two negative minima at about 210 and 225 nm appear along with a maximum at 195 nm. The curve crosses the baseline from negative to positive at about 202 nm. The amplitudes of positive and negative CD bands increased in the presence of 40% TFE. The CD patterns observed in the presence of 20-40% TFE are reminiscent of an  $\alpha$  helix CD spectrum, although the intensities of positive and negative bands are reduced as compared with the reference CD spectrum for 100%  $\alpha$  helix. This means that other secondary structures are also present under these conditions. In Figure 3 (insert) the mean residue molar ellipticity of peptide II at 222 nm,  $[\Theta]_{222}$ , is plotted against the concentration of TFE in solution. One can see that  $[\Theta]_{222}$  increases with increasing TFE concentration. At 50% TFE a plateau is reached. To estimate the helix content of peptide II at various concentrations of TFE we have used two methods. The content of  $\alpha$  helix,  $f_h$ , can be approximately calculated as:

$$f_h = \frac{[\Theta]_{222} - [\Theta]_{222}^c}{[\Theta]_{222}^h (1 - \chi/n) - [\Theta]_{222}^c} \quad [1]$$

where

$[\Theta]_{222}^c$  is the mean residue molar ellipticity for 0% helix.



**Figure 3:** CD spectra of peptide II in aqueous buffer in the presence and absence of TFE. 1, 0% (v/v) TFE; 2, 20% TFE; 3, 40% TFE.  $[\Theta]$  is the mean residue molar ellipticity expressed in  $\text{deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$ . Insert: A plot of  $[\Theta]_{222}$  against the concentration of TFE in solution. CD spectra were recorded in 1 mM sodium cacodylate buffer (pH 7.0) in the presence of 0.1 M  $\text{NH}_4\text{F}$  and various concentrations of TFE. Concentration of the peptide was  $1 \cdot 10^{-5}$  M.

$[\Theta]_{222}^h$  is the reference value of molar ellipticity for 100% helix.  $n$  is the number of residues per helix.  
 $\chi$  is the chain length-dependent factor.

Assuming that helical regions in each of the four arms of peptide II are confined to the same residues as in native 434 *cro* (24), we set  $n=9$  into Equation [1]. Chang *et al.* have found that  $\chi=2.5$  for  $n=9$  (23). Unordered peptides usually have an intense CD band near 200 nm and a shoulder near 220 nm. However, ellipticity at 222 nm is a subject of considerable variance:  $[\Theta]_{222}^c \approx 3.9 \cdot 10^3 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$  for poly-L-lysine at pH 5.7 (25), but it takes a negative value of about  $-1.8 \cdot 10^3 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$  for (Pro-Lys-Leu-Lys-Leu) $_n$  in salt-free water (26). Goodman and Kim have obtained a value of about  $890 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$  for ellipticity  $[\Theta]_{222}^c$  of a synthetic peptide containing carboxyterminal twelve residues of bovine trypsin inhibitor (27). Using this value for  $[\Theta]_{222}^c$ , one can calculate that helix content of peptide II in the presence of 20% TFE is about 27%. It increases up to 47% in the presence of 40% TFE and approaches 52% at 50-70% TFE. This indicates that helix formation by the peptide is not complete



at room temperature: 12 out of 23 residues in each arm of the peptide are in the  $\alpha$ -helical conformation. As a test for aggregation, CD spectra were taken at a series of peptide concentrations (2-10  $\mu\text{M}$ ) in the presence of 20% TFE. No changes in molar ellipticity at 222 nm were observed as the peptide was diluted. This suggests that helix formation by peptide II is a monomolecular process.

In our previous studies of peptide I (8), we have used CD spectra for poly-L-lysine in  $\alpha$ -,  $\beta$ - and unordered conformations to estimate contents of various secondary structures in the peptide I in the presence of differing TFE concentrations. The content of  $\alpha$ -,  $\beta$ - and unordered conformations, estimated as outlined by Stokrova *et al.* (28), was found to be 0.16, 0.21 and 0.63, respectively, in the presence of 20% TFE and to be 0.40, 0.20 and 0.40, respectively, in the presence of 40% TFE (8). Since the mean residue molar ellipticities at 195 and 222 nm for peptides I and II were nearly identical, these estimates can be extended to the peptide II.

Comparing the results of component analysis with the estimates obtained from Equation [1] we conclude that Equation [1] overestimates the actual helix content in the peptide. Similar analysis confirms the absence of  $\alpha$ -helical conformation in the peptide II when no TFE was present in solution. In this case the measured CD spectra can be well represented as linear combinations of spectra characteristic of  $\beta$ -sheet and random-coiled conformations. Using a value of  $[\Theta]_{222} = -1.3 \cdot 10^4 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$  for the molar ellipticity of poly-L-lysine in  $\beta$ -conformation (25), the content of  $\beta$  strand was estimated to be about 24% at 0% TFE, provided that  $[\Theta]_{222}^c = 890 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$  (27). The uncertainty in this estimate results primarily from the fact that reference spectra for  $\beta$ -sheet and unordered conformations are subjects of considerable variance (29).

#### CD Spectra of Complexes between Peptide II and Various Nucleic Acids

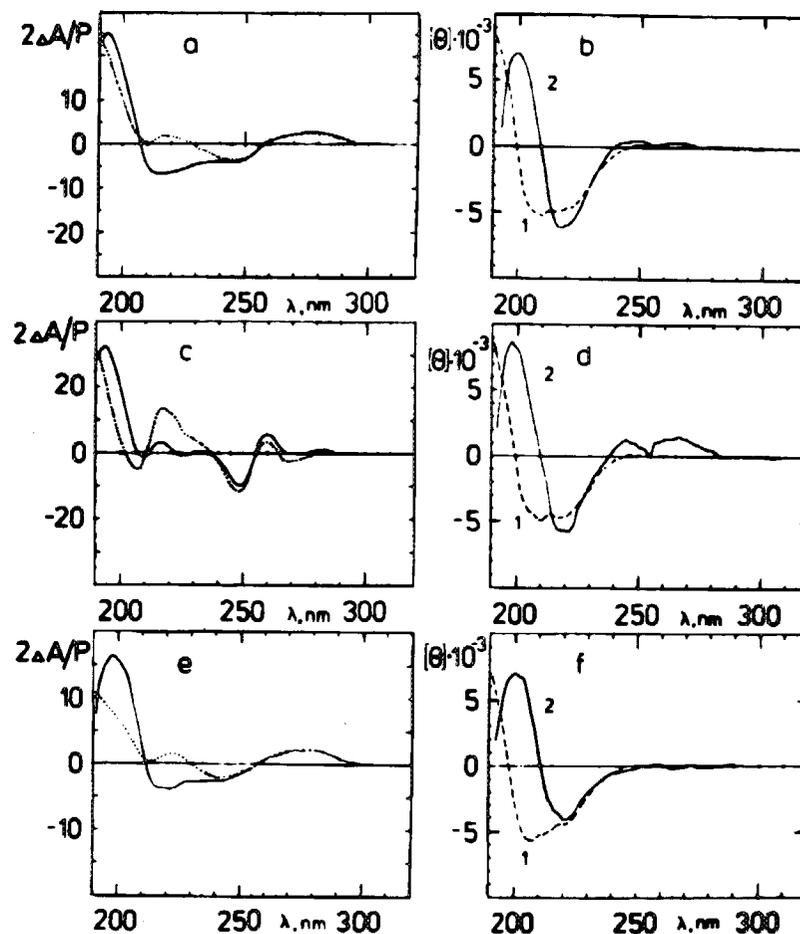
Upon mixing peptide II and calf thymus DNA in the presence of 20% TFE the CD spectra are dramatically changed in the wavelength region 190-230 nm, where the free peptide has intensive positive and negative CD bands (Figure 4). Intrinsic CD of naked DNA at wavelengths 200-230 nm is small, as compared with the CD signal from the free peptide under conditions used in our experiment. Since the free peptide has a negligible intrinsic CD at wavelengths above 250 nm, the CD characteristics of the long-wavelength positive DNA band can be used to monitor the state of DNA in the peptide-DNA complexes. From literature data (30) it is known that CD characteristics of this band are very sensitive to small changes of the winding angle per base pair for a nucleic acid. On adding peptide II to DNA only minor changes are observed in the wavelength region 250-300 nm (Figure 4). This indicates that DNA does not undergo considerable structural changes (such as a transition from the B to an A-helical form) upon complex formation with the peptide. The CD spectrum of a peptide-DNA mixture cannot be represented as a sum of spectra of the free peptide and naked DNA. The difference CD spectrum obtained by subtracting the spectrum of DNA from the spectrum of a peptide-DNA mixture has positive and negative CD bands at 195 and 220 nm, respectively (Figure 4). Similar CD spectral changes are observed upon mixing of peptide II with DNA in the presence of 40% TFE (9). Since

TFE. 1.0% (v/v) TFE;  $\text{cm}^2 \cdot \text{dmol}^{-1}$ . Insert: A  $\text{dm}^2$  in 1 mM sodium  $\text{dm}^2$  of TFE. Concentration

ix. n is the

re confined to the Chang *et al.* have intense CD band 22 nm is a subject ly-L-lysine at pH  $\text{dmol}^{-1}$  for (Pro- obtained a value peptide containing Using this value presence of 20% and approaches de is not complete

difference CD spectra look similar to the reference CD spectrum for  $\beta$ -sheet conformation (25,26,29), we suggest that conversion of the  $\alpha$  helical structure to a  $\beta$ -sheet conformation takes place upon binding of the peptide to DNA. Figure 4 shows that similar CD changes are observed upon mixing of peptide II with poly(dA) · poly(dT),



**Figure 4:** Figure 4: Changes observed in CD spectra of peptide II upon binding to calf thymus DNA and synthetic polynucleotides. CD spectra are shown for calf thymus DNA (a), poly(dA) · poly(dT) (c) and poly[d(AC)] · poly[d(GT)] (e) in the absence and presence of the peptide. The CD spectrum of a nucleic acid alone and of a mixture of the nucleic acid with the peptide are shown by dotted and solid lines, respectively.  $2\Delta A/P$  is the measured dichroism calculated per mole of DNA base pairs and 1 cm pathlength cell. (b, d, f), difference CD spectrum obtained by subtracting the CD spectrum of the corresponding nucleic acid from the spectrum of the peptide-DNA mixture (2) compared with the CD spectrum of the free peptide (1). Difference CD spectra for peptide II complexed with calf thymus DNA (b), poly(dA) · poly(dT) (d) and poly[d(A-C)] · poly[d(G-T)] (f) are shown.  $[\theta]$  is the mean residue molar ellipticity expressed in  $\text{deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$ . Concentrations of DNA, poly(dA) · poly(dT) and poly[d(A-C)] · poly[d(G-T)] were equal to  $1.3 \cdot 10^{-4}$  M,  $1.0 \cdot 10^{-4}$  M and  $1.2 \cdot 10^{-4}$  M (base pairs), respectively. Concentration of the peptide was  $6.1 \cdot 10^{-6}$  M (a,b,c,d) and  $5.1 \cdot 10^{-6}$  M (e,f). CD spectra were recorded at  $20^\circ\text{C}$  in 1 mM sodium cacodylate buffer (pH 7) containing 0.1 M  $\text{NH}_4\text{F}$  and 20% TFE.

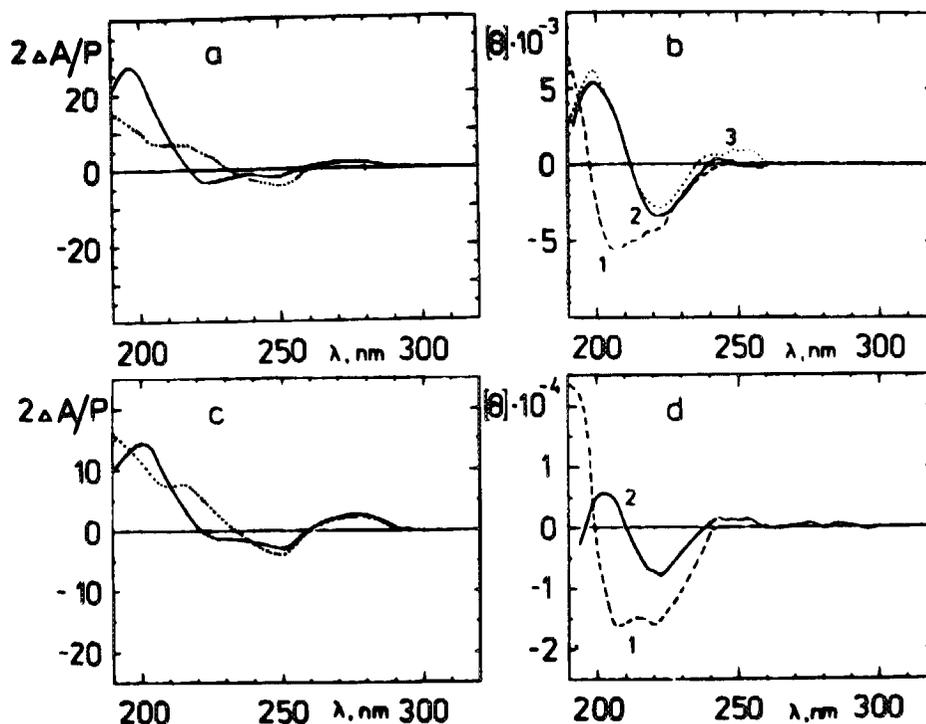
β-sheet conformation to a β-sheet structure. Figure 4 shows that poly(dA) · poly(dT),



Figure 4 shows the CD spectra of calf thymus DNA and poly(dA) · poly(dT) (c) and poly(dG) · poly(dC) (a) and poly(dG) · poly(dC) (b) in a 10 mm path length cell. (b, d, f), CD spectra of the free peptide poly(dA) · poly(dT) (d) and poly(dG) · poly(dC) (f) in the presence of the peptide poly(dA) · poly(dT) (e) and poly(dG) · poly(dC) (g) in the presence of sodium cacodylate.

poly[d(AC)] · poly[d(GT)] in the presence of 20% TFE. The difference CD patterns generated for these systems exhibit a remarkable similarity with each other. However, CD spectral profiles of naked calf thymus DNA and poly(dA) · poly(dT) are markedly different, reflecting differences in their nucleotide composition and conformation (31). We interpret these observations as indicating that bound peptide molecules share a conformation feature in common in all these complexes and that structure of nucleic acid is not markedly perturbed upon complex formation with the peptide.

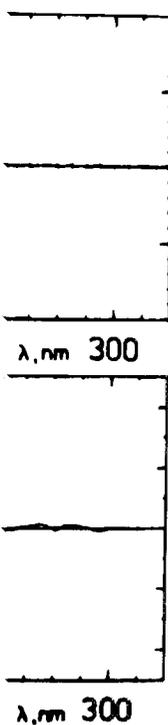
It is of interest to compare the magnitudes of CD extrema in the difference spectra for peptide II with those of the reference CD spectrum for a β-sheet conformation (25). Care is needed in making such a comparison since the equilibrium between different conformation states of the peptide appear to be altered in the presence of DNA. One should either estimate the amount of bound peptide in the system under study or deal with the experimental situation when DNA was taken in a large molar excess over the peptide. In the latter case one could neglect a contribution of the free peptide into an overall CD signal. Unfortunately, at DNA concentration levels on the order of  $1 \cdot 10^{-3}$  M (base pairs) the binding reaction is accompanied by aggregation, as revealed from absorbancy reading at 320 nm and sedimentation experiments. Aggregation was negligible in the presence of 10 times lower DNA concentration, provided that molar peptide to DNA base pair ratio was less than 0.1. Our experiments were made in this range of DNA concentrations at peptide/DNA base pair ratio of about 0.05. The apparent absorbancy at 320 nm for complexes of peptide II with poly(dA) · poly(dT) was less than 0.02. From competition-type experiments made in the presence of minor groove-binding antibiotic distamycin A we estimated that binding constant of peptide II to poly(dA) · poly(dT) is greater than  $10^5$  M<sup>-1</sup>. This means that more than 90% of the peptide molecules were bound to poly(dA) · poly(dT) under conditions used in our experiments. From difference CD spectrum for peptide II complexed with poly(dA) · poly(dT) (Figure 4) we have found that  $[\Theta]_{196} = 1.1 \cdot 10^4$ ,  $[\Theta]_{220} = -7.4 \cdot 10^3$  and  $[\Theta]_{196}/[\Theta]_{220} = -1.5$ . The ellipticities are expressed in deg · cm<sup>2</sup> · dmol<sup>-1</sup>. Nearly the same numerical values of  $[\Theta]_{196}$  and  $[\Theta]_{220}$  were obtained for complexes of peptide II with calf thymus DNA and poly[d(AC)] · poly[d(GT)] in the presence of 20% TFE (Figure 4) and for complexes of peptide I with poly(dA) · poly(dT) (8). For poly-L-lysine in β-conformation  $[\Theta]_{196} = 3.2 \cdot 10^4$ ,  $[\Theta]_{220} = -1.6 \cdot 10^4$  and  $[\Theta]_{196}/[\Theta]_{220} = -1.8$  (25). This indicates that both the intensity ratio and shape of difference CD spectrum for peptide II are reminiscent of a β-sheet CD pattern. However, there are quantitative differences in ellipticities at 196 and 220 nm. For example, the positive 196 nm band in the difference spectrum is only about half as strong as it is in the CD spectrum for β conformation (25). The observed difference in magnitude of CD extrema could be attributed to an environment effect of a bound nucleic acid on the rotational strengths of  $n\pi^*$  and  $\pi\pi^*$  transitions. In this respect it should be noted that  $[\Theta]_{196}$  and  $[\Theta]_{220}$  values of sodium dodecylsulfate - poly-L-lysine complex were smaller than corresponding values for poly-L-lysine in the β conformation induced in alkaline solution at 50°C. This can be attributed to the environment effect of bound surfactant molecules on the rotatory strengths of the  $n\pi^*$  transition and  $\pi\pi^*$  transitions (32,33). The hydrophobic interactions between bound surfactant molecules are known to be the stabilizing factor for β-sheet conformation (34).



**Figure 5:** Changes observed in CD spectra of peptides II and III upon their binding to synthetic DNA oligomers  $O_I$  and  $O_{II}$ . a) CD spectrum of the operator  $O_I$  alone (---) and of a mixture of peptide II with the operator fragment (---). A similar CD pattern was obtained for a mixture of peptide II with the DNA oligomer  $O_{II}$  not shown. b) CD spectrum of free peptide II (1) and difference CD spectra (2,3) obtained by subtracting the CD spectra of DNA oligomers  $O_I$  and  $O_{II}$  from the spectra of corresponding complexes with the peptide. Difference CD spectra for peptide II bound to the DNA oligomers  $O_I$  and  $O_{II}$  are shown by dotted and solid lines, respectively. c) CD spectrum of the DNA oligomer  $O_{II}$  alone (---) and of a mixture of peptide III with the DNA oligomer (---). d) CD spectrum of free peptide III (1) and difference CD spectrum for peptide III bound to the DNA oligomer  $O_{II}$  (2). Concentrations of the oligomers  $O_I$  and  $O_{II}$  were  $6.7 \cdot 10^{-5}$  M and  $1.06 \cdot 10^{-4}$  M (base pairs), respectively. Concentrations of peptide II in complexes with DNA oligomers  $O_I$  and  $O_{II}$  were  $5.2 \cdot 10^{-6}$  M and  $9.1 \cdot 10^{-6}$  M, respectively. Concentration of peptide III was  $8.3 \cdot 10^{-6}$  M. Experimental conditions were identical to those in Figure 4.

#### *Binding of Peptides II and III to Synthetic Operators $O_I$ and $O_{II}$*

CD changes observed upon binding of peptide II to DNA oligomers  $O_I$  and  $O_{II}$  are similar to those described above for binding of the peptide to calf thymus DNA and synthetic polynucleotides (see Figures 4 and 5). This suggests that both the specific and nonspecific binding reactions are accompanied by helix- $\beta$  transition in the DNA binding domains of the peptide. The difference CD spectra generated for peptide II in the presence of duplexes  $O_I$  and  $O_{II}$  were almost identical. Since peptide II is an assembly of four linear peptide chains, each corresponding to a modified helix-turn-helix motif of the 434 *cro*, a possibility exists that  $\beta$ -structure is stabilized by interactions between adjacent chains in the peptide molecule. To test the validity of



g to synthetic DNA  
l of a mixture of  
s obtained for a  
um of free peptide  
A oligomers O<sub>I</sub> and  
pectra for peptide II  
ely. c) CD spectrum  
ligomer (---). d) CD  
to the DNA oligomer  
(base pairs), respec-  
ere 5.2 · 10<sup>-6</sup> M and  
tal conditions were

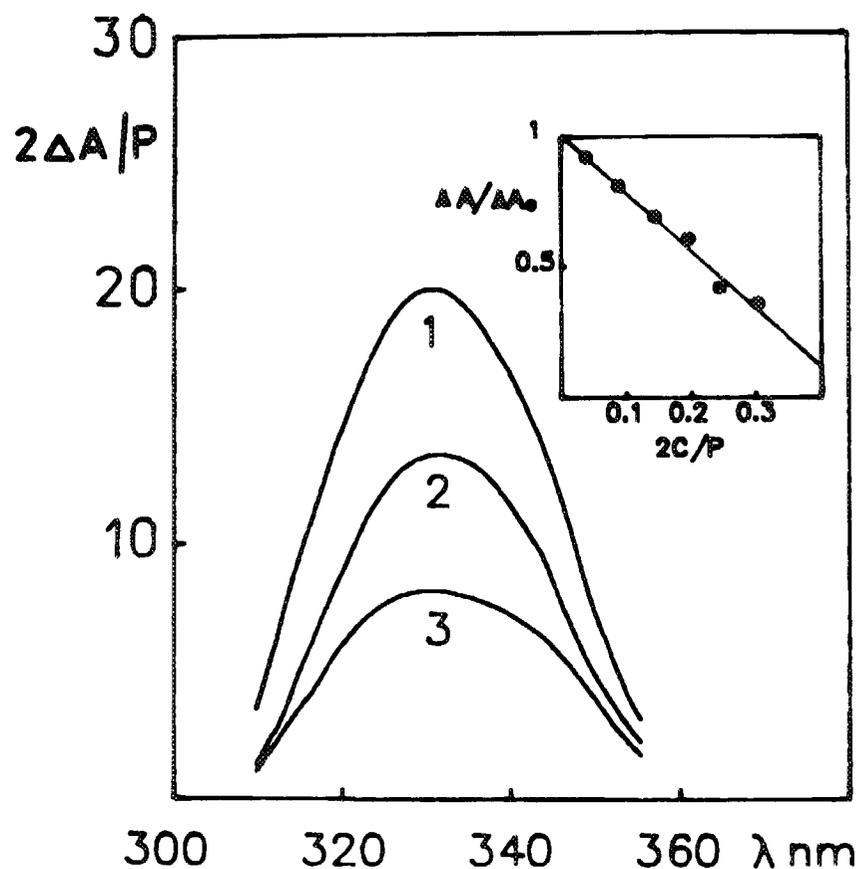
s O<sub>I</sub> and O<sub>II</sub> are  
ymus DNA and  
both the specific  
ransition in the  
rated for peptide  
peptide II is an  
modified helix-  
is stabilized by  
st the validity of

this hypothesis we examined the interaction with duplexes O<sub>I</sub> and O<sub>II</sub> of a linear peptide III containing the modified helix-turn-helix motif of the 434 *cro* repressor (Figure 1). The CD spectrum of peptide III in aqueous buffer containing 20% TFE shows double minima at 210 and 220 nm and a maximum near 190 nm that are typical of  $\alpha$ -helical conformation. The estimated helix content in 20% TFE was about 60%. It is greater than that found for peptide II under the same conditions. This result can be explained, if a part of the twelve N-terminal residues in the sequence of peptide III, not found in the peptide II, are in the  $\alpha$ -helical conformation. Consistent with this are X-ray data showing that residues 7-14 of the 434 *cro* exist in the  $\alpha$  helix conformation (24). In the difference CD spectrum obtained by subtracting the CD spectrum of duplex O<sub>II</sub> from the spectrum of mixture of the peptide with the O<sub>II</sub> the double minima at about 210 and 225 nm disappeared and the CD signal at 190 nm decreased, thus giving rise to a CD spectral profile which seems to be more similar to a  $\beta$ -like CD pattern (Figure 5).

*Displacements of Distamycin A from Poly(dA) · Poly(dT) and Synthetic Operators by Peptides II and III*

The two related antibiotics, distamycin A and netropsin, have received much attention in recent years as examples of sequence-specific minor groove binding molecules. The recently published crystal structures of DNA complexes with netropsin (35) and distamycin A (36) confirm earlier proposals that these two antibiotics bind in the minor DNA groove and interact specifically with AT- base pairs (37,38). The NH groups of the four distamycin amide groups form hydrogen bonds with the adenine N3 and thymine O2 atoms, and the aromatic pyrrole rings forms van der Waals contacts with the AT- base pairs.

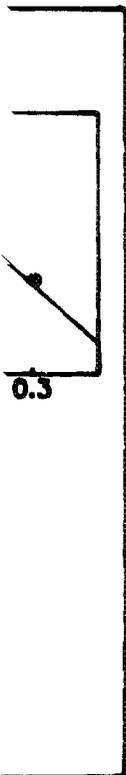
In the present section, we report data showing that there is a competition between distamycin A and peptide II for binding to poly(dA) · poly(dT) and synthetic operator fragment. The binding of distamycin A to DNA can be monitored by CD spectroscopy in the wavelength region 310-360 nm. Since the free antibiotic is optically inactive, CD measurements provide a rather simple and sensitive tool for studies of binding equilibria between distamycin A and nucleic acids. The CD amplitude at 320 nm is known to be proportional to the concentration of the antibiotic bound to a nucleic acid. On adding peptide II to poly(dA) · poly(dT) preincubated with distamycin A, the CD amplitude at 320 nm is found to decrease progressively with increasing concentration of the peptide (Figure 6). The experimentally measured value for saturation level of binding of distamycin A to poly(dA) · poly(dT) has been found to be one bound antibiotic molecule per five base pairs. Since no precipitation of poly(dA) · poly(dT) was observed on adding peptide II, we interpret these observations as indicating that peptide II displaces distamycin A from poly(dA) · poly(dT). About half of the bound antibiotic molecules dissociate from poly(dA) · poly(dT) at molar peptide to DNA base pair ratio of about 0.2 (Figure 6). This means that the apparent binding constant of peptide II to poly(dA) · poly(dT) is comparable with that of distamycin A. The binding constant of distamycin A to poly(dA) · poly(dT) in the presence of 0.1 M NaCl and 20% TFE has been estimated to be greater than 1 · 10<sup>6</sup> M<sup>-1</sup>.



**Figure 6:** Titrations of distamycin-poly(dA) · poly(dT) complexes with peptide II. CD spectra of complexes between distamycin A and poly(dA) · poly(dT) were recorded in the absence and presence of peptide II. Concentrations of the peptide were as follows: 1. 0; 2.  $3.3 \cdot 10^{-6}$ ; 3.  $5.2 \cdot 10^{-6}$  M. Concentration of poly(dA) · poly(dT) was  $2.2 \cdot 10^{-5}$  M (base pairs). Concentration of distamycin A was  $8.2 \cdot 10^{-6}$  M. Insert: A plot of the normalized CD signal at 320 nm,  $(\Delta A/\Delta A_0)$ , against  $2C/P$ , the molar ratio of peptide II to DNA base pairs.  $\Delta A_0$  and  $\Delta A$  are the measured CD amplitudes per 1 cm pathlength cell for complexes of distamycin A with poly(dA) · poly(dT) in the absence and presence of the peptide, respectively.  $2\Delta A/P$  is the measured CD amplitude calculated per mole of DNA base pairs and 1 cm pathlength cell. Titrations were made at 20°C in 1mM sodium cacodylate buffer (pH 7.0) containing 0.1 M NaCl and 20% TFE.

Titrations with the peptide II of the oligonucleotide duplex  $O_I$  complexed with the antibiotic yield similar results (Figure 7). Again, the binding of peptide II to the oligonucleotide is accompanied by displacement of distamycin A from the operator  $O_I$ . Figure 8 shows typical Scatchard isotherms obtained for binding of distamycin A to the synthetic operator  $O_I$  in the absence and presence of peptide II. Similar data were obtained for binding of distamycin A to a synthetic oligonucleotide duplex  $O_{II}$ . It is well-known that intercept of a binding isotherm on the vertical axis,

$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} r/m,$$



$\lambda$  nm

CD spectra of com- and presence of pep- M. Concentration of s  $8.2 \cdot 10^{-6}$  M. Insert: ratio of peptide II to cell for complexes of s respectively.  $2\Delta A/P$  is length cell. Titrations NaCl and 20% TFE.

plexed with the peptide II to the om the operator ng of distamycin e II. Similar data ptide duplex  $O_{II}$ - al axis,

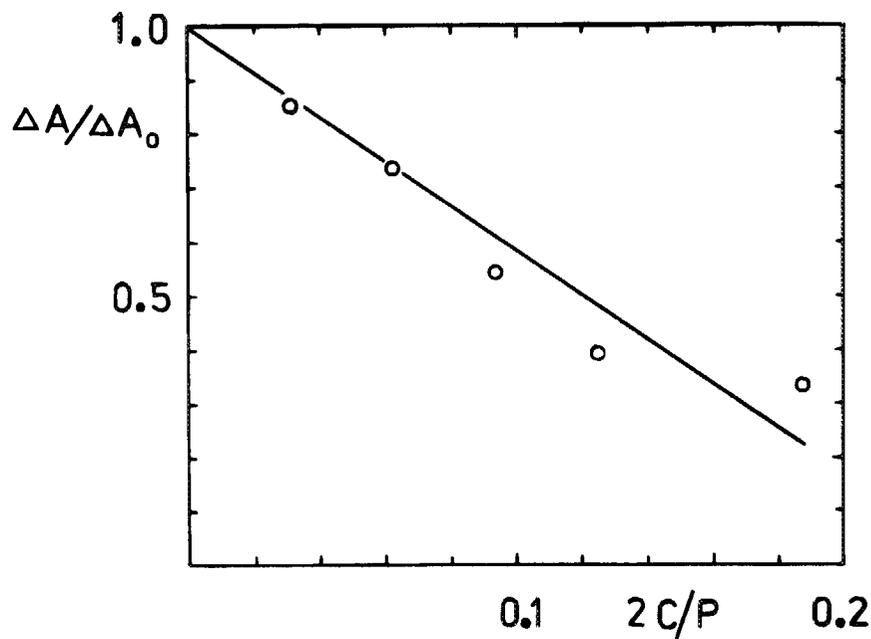


Figure 7: Displacement of distamycin A from the synthetic operator  $O_1$  by peptide II. Concentration of the DNA oligomer was  $4.2 \cdot 10^{-6}$  M. Concentration of distamycin A was  $3 \cdot 10^{-6}$  M. Experimental conditions were identical to those in Figure 6.

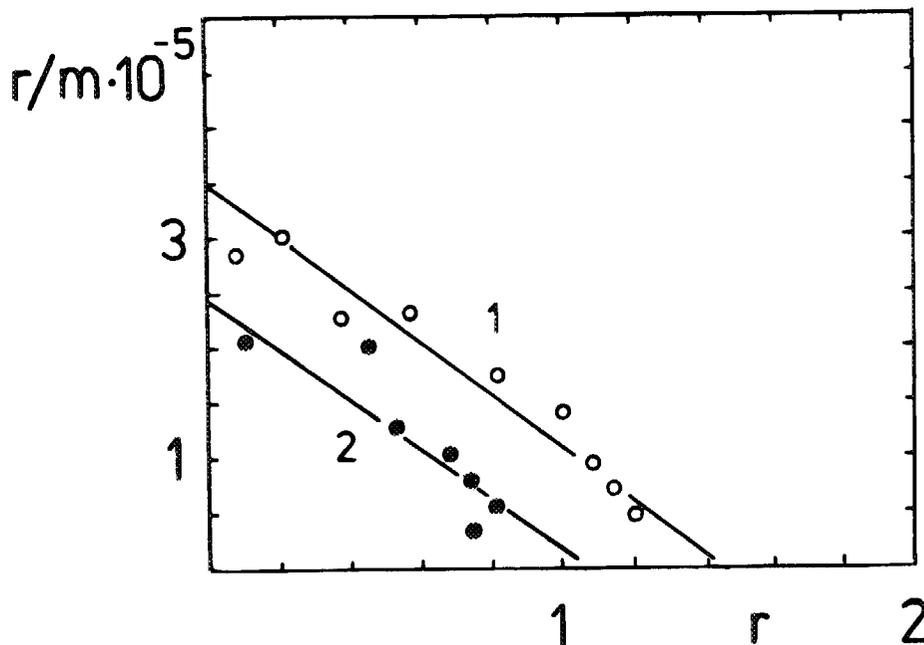


Figure 8: Binding isotherms of distamycin A to the DNA oligomer  $O_1$  in the absence (curve 1) and presence (curve 2) of peptide II. Concentration of the oligomer  $O_1$  was  $4.74 \cdot 10^{-6}$  M. Concentration of the peptide was  $4.9 \cdot 10^{-6}$  M. Experimental conditions were identical to those in Figure 6.

is equal to the average binding constant of a ligand to an isolated site on a polynucleotide lattice (39,40). Here  $r$  is the ratio of moles of bound distamycin A to moles of the oligonucleotide duplex.  $m$  is the concentration of the free antibiotic. We shall denote by

$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} (r/m)_E$$

and

$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} (r/m)_R$$

the intercepts on the vertical axis of the experimental isotherms of distamycin A to the DNA oligomer measured in the presence and absence of added peptide, respectively. The binding isotherm obtained in the absence of peptide will be referred to as the reference binding isotherm. From Figure 8 one can conclude that the apparent number of binding sites,  $n_{app}$ , for the antibiotic on the DNA oligomer  $O_1$  decreases in the presence of peptide II. This indicates that peptide II occupies a part of potential binding sites for distamycin A on the operator. Similar results were obtained from titrations with distamycin A of a synthetic operator  $O_{R3}$  preincubated with  $\lambda$  phage *cro* repressor. From CD studies and gel retardation experiments it was found that distamycin A displaced *cro* repressor from the operator DNA (41). If simultaneous binding of distamycin A and peptide to the same DNA oligomer is not possible, each oligonucleotide duplex may exist in one of the three states: it can be either free or complexed with distamycin A and peptide. In this case the following expressions can be derived for the intercepts on the vertical axis of binding isotherms obtained in the presence and absence of added peptide.

$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} (r/m) = \langle K \rangle / Z \quad [2]$$

$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} (r/m) = \langle K \rangle \quad [3]$$

where

$$\langle K \rangle = N^{-1} \sum_{j=1}^{N-L+1} K_j \quad [4]$$

$$Z = 1 + K_1 C_1 \quad [5]$$

ed site on a poly-  
amycin A to moles  
ntibiotic. We shall

Here  $\langle K \rangle$  is the average binding constant of distamycin A to the oligonucleotide duplex containing  $N$  base pairs.  $K_j$  is the binding constant of distamycin A to the  $j$ -th site on the oligonucleotide duplex ( $1 \leq j \leq N-L+1$ ).  $L$  is the size of binding site for distamycin A on DNA.  $Z$  is the partition function for the thermodynamic system comprising oligonucleotide duplexes and peptide molecules only.  $K_1$  is the binding constant of the peptide to the DNA oligomer.  $C_1$  is the concentration of the free peptide. It is assumed that only one peptide molecule can be bound to each DNA oligomer. This seems to be a reasonable assumption, since a pair of helix-turn-helix motifs of the peptide can occupy simultaneously the two outer symmetry-related regions on the operator DNA. In the limit, when  $r \rightarrow 0$ , the average occupancy of the oligonucleotide duplex by bound peptide,  $\eta$ , is given by the following standard relation:

$$\eta = \frac{\partial \ln Z}{\partial \ln C_1} \quad [6]$$

where

$$C_1 = C - \eta[O] \quad [7]$$

Here  $C$  is the peptide concentration in the system under study.  $[O]$  is the molar concentration of the synthetic DNA oligomer. In the absence of competing ligand the fraction of the free oligonucleotide is given by

$$[O]_f/[O] = 1/Z \quad [8]$$

where  $[O]_f$  is the concentration of the free oligonucleotide. From experimental curves shown in Figure 8 it can be found that  $\langle K \rangle = (3.5 \pm 0.2) \cdot 10^5 \text{ M}^{-1}$ ,  $K_1 C_1 = 0.46 \pm 0.20$ ,  $\eta = 0.32 \pm 0.09$  and  $[O]_f/[O] = 1 - \eta = 0.68 \pm 0.09$ . Inserting numerical values for  $C$  and  $[O]$  into Equation [7] one can find that  $K_1 = (1.4 \pm 0.4) \cdot 10^5 \text{ M}^{-1}$ . The fraction of free DNA oligomers  $[O]_f/[O]$  can be roughly estimated from Equations [9] and [10]:

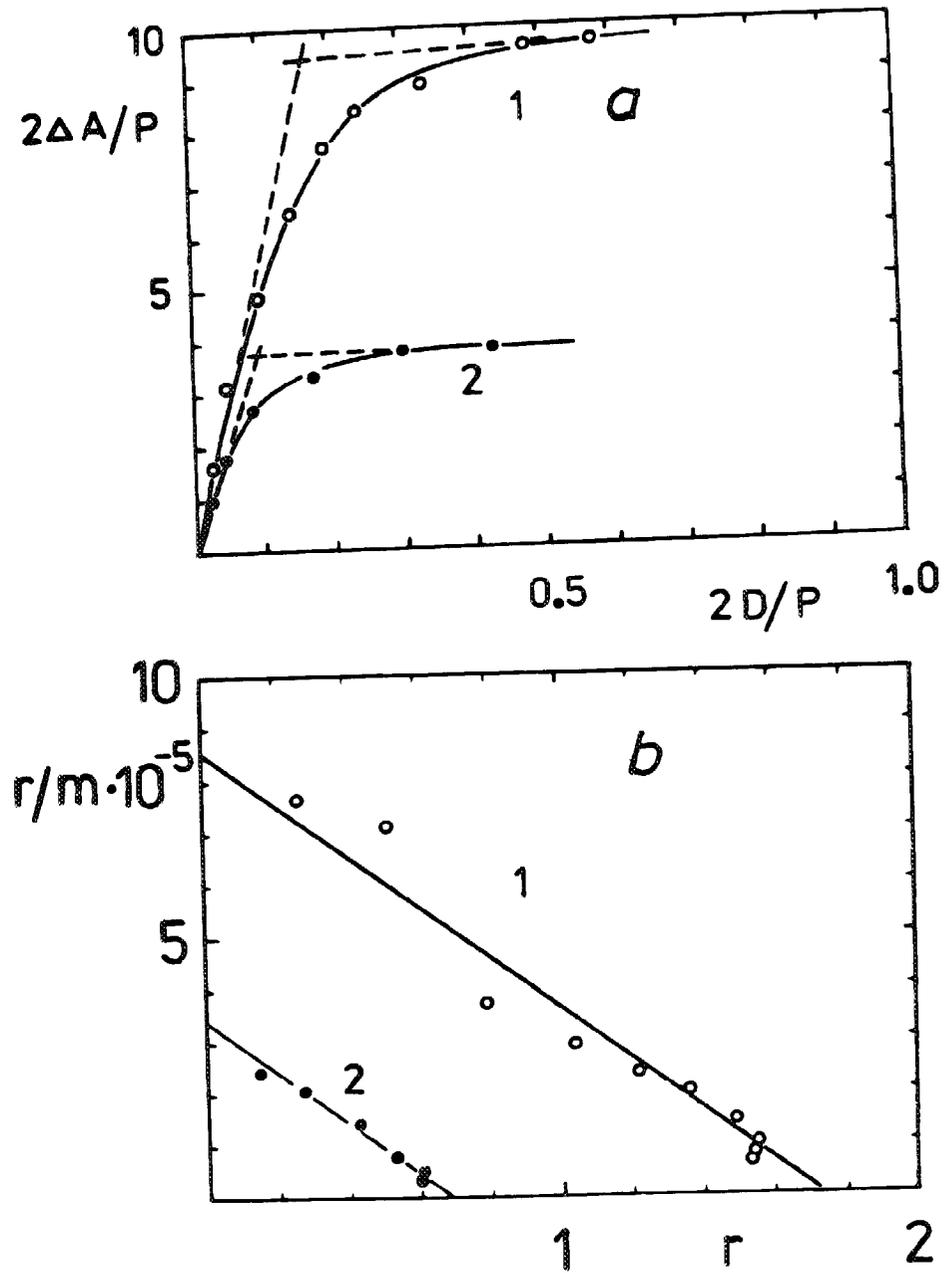
$$[2] \quad [O]_f/[O] = (n_{app})_E / (n_{app})_R \quad [9]$$

$$[O]_f/[O] = (r_{max})_E / (r_{max})_R \quad [10]$$

[3] where  $(n_{app})_R$  and  $(n_{app})_E$  are the apparent numbers of binding sites for distamycin A determined from titrations of naked DNA oligomers and DNA oligomers partially complexed with the peptide.  $(r_{max})_R$  and  $(r_{max})_E$  are the saturation levels of binding of distamycin A to the oligonucleotide duplex in the absence and presence of peptide, respectively.

[4] From the intercepts of the experimental isotherms on the horizontal axis (see Figure 8) one can find that  $(n_{app})_R = 1.42 \pm 0.08$ ,  $(n_{app})_E = 1.05 \pm 0.08$  and  $[O]_f/[O] = 0.74 \pm 0.10$  at peptide/DNA oligomer ratio of about 1. From literature data (14) it is known that distamycin A interacts preferentially with runs of AT-pairs and avoids regions containing GC-pairs. It covers five DNA base pairs upon binding. Inspection of the

[5]



**Figure 9:** (a) Titration with distamycin A of the free oligonucleotide  $O_{II}$  (curve 1) and oligonucleotide in the presence of peptide III (curve 2). Concentration of the DNA oligomer was  $3.3 \cdot 10^{-6}$  M. Concentration of the peptide was  $3.2 \cdot 10^{-6}$  M.  $2D/P$  is the molar ratio of added distamycin A to the DNA oligomer. (b) Binding isotherms of distamycin A to the DNA oligomer  $O_{II}$  in the absence (curve 1) and presence (curve 2) of peptide III. Experimental conditions were identical to those in Figure 6.

nucleotide sequence for the oligonucleotide duplex  $O_I$  (Figure 2) shows that it contains 10 successive AT-base pairs to which either one or two distamycin molecules could be bound. This is in agreement with the experimental value for the apparent number of binding sites per oligonucleotide duplex ( $n_{app} \approx 1.4$ ).

Figure 9 shows typical CD titration curves and Scatchard isotherms obtained for binding of distamycin A to the oligonucleotide duplex  $O_{II}$  in the absence and presence of peptide III. These experiments demonstrate that binding equilibrium between distamycin A and the oligonucleotide is greatly affected by the peptide. Both

$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} r/m$$

and  $n_{app}$  are reduced more than two-fold at peptide/DNA oligomer ratio of about 1. Similarly, the saturation level of binding of distamycin A to the oligomer  $O_{II}$  preincubated with the peptide was about 2.5 times lower than that found from the binding of distamycin A to the free oligonucleotide. This means that in the presence of peptide III a part of potential binding sites for distamycin A on the  $O_{II}$  (or a part of DNA duplexes) become inaccessible to further binding events. To estimate the binding constant of the peptide to the  $O_{II}$  one should take into account that peptide III, unlike the peptide II, has only one helix-turn-helix motif. As a consequence, it can form complexes with a stoichiometry of either one or two monomers per oligonucleotide duplex. In this case

$$Z = 1 + 2K_1C_1 + K_1^2C_1^2 \quad [11]$$

As before, we neglect situations when both distamycin A and peptide bind simultaneously to the same DNA fragment. In this case Equations [2],[3], [6],[8]) and [11] can be used to estimate  $K_1C_1$ ,  $\eta$  and  $[O]_f/[O]$  from the experimental values of

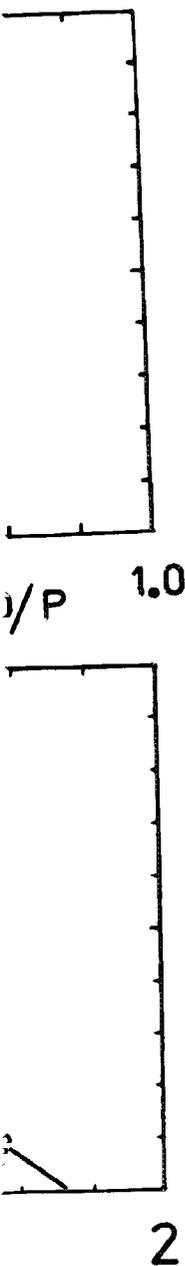
$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} (r/m)_E = (8.5 \pm 0.5) 10^5 M^{-1}$$

and

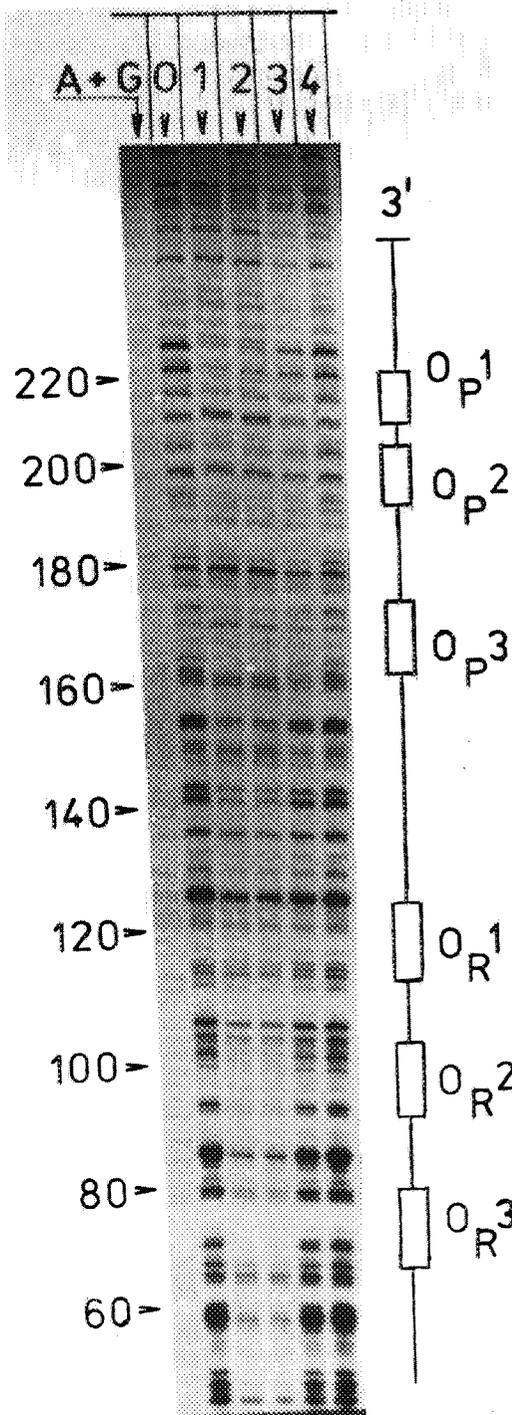
$$\lim_{\substack{r \rightarrow 0 \\ m \rightarrow 0}} (r/m)_R = (3.4 \pm 0.1) 10^5 M^{-1}$$

We have found that  $K_1C_1 = 0.58 \pm 0.06$ ,  $\eta = 0.73 \pm 0.10$  and  $[O]_f/[O] = 0.40 \pm 0.03$ . Inserting  $C = 3.2 \cdot 10^{-6} M$  and  $[O] = 3.36 \cdot 10^{-6} M$  into Equation [7] and using the estimated value for  $K_1C_1$ , one can find that  $5 \cdot 10^5 < K_1 < 1.6 \cdot 10^6 M^{-1}$ .

The fraction of DNA oligomers inaccessible for the binding to distamycin A can be estimated from a comparison of CD amplitudes at 320 nm measured at saturation



(a) and oligonucleotide in  $10^{-6} M$ . Concentration of the DNA oligomer. (b) (curve 1) and presence (curve 2)



**Figure 10:** Digestion of the DNA fragment A by the acidic DNase in the presence and absence of peptide II. The  $^{32}\text{P}$  end-labeled DNA fragment A containing operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  and unlabeled plasmid pUC9 were incubated with the peptide for 20 min and then digested by the enzyme. The numbers on the left-hand side of the gel correspond to those of the DNA sequence (see Figure 12). The DNA fragment was labeled selectively at the 3'-end of the bottom strand with [ $\alpha$ - $^{32}\text{P}$ ] dATP and Klenow enzyme. Concentration of the carrier DNA was  $4 \cdot 10^{-6}$  M (base pairs). Concentrations of the peptide were as follows: lane 1,  $2.3 \cdot 10^{-6}$  M; lane 2,  $1.2 \cdot 10^{-6}$  M; lane 3,  $6 \cdot 10^{-7}$  M; lane 4,  $3 \cdot 10^{-7}$  M. Lane 0, control, i.e., digestion of the free DNA. Concentrations of the enzyme used for digestion of naked DNA and DNA complexes with the peptide were  $0.5 \mu\text{g/ml}$  and  $1-2 \mu\text{g/ml}$ , respectively. A+G, positions of purines in the sequence cut by formic acid/diphenylamine. The operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  and some pseudooperator sites serving as interaction sites for peptide II are marked on the right-hand side of the gel.

levels of binding for complexes of distamycin A with naked DNA oligomers and DNA oligomers partially complexed with the peptide. As can be seen from Figure 9, at molar peptide/DNA oligomer ratio of about 1, the fraction of DNA inaccessible for binding of distamycin A is equal to about 60%, in agreement with the estimate obtained earlier from statistical mechanical considerations. This indicates that the proposed model gives self-consistent values for binding parameters.

#### *Sequence-Selective Binding of Peptide II to DNA*

To determine whether peptide II is capable of recognizing any particular sequences on natural DNA fragments of known nucleotide sequence footprinting experiments were performed using either bovine pancreatic DNase I or acidic DNase isolated from walleye pollock liver (20). Figure 10 shows the peptide II footprint pattern for the bot-

the DNA fragment A by presence and absence of labeled DNA fragment A. O<sub>R</sub>1, O<sub>R</sub>2 and O<sub>R</sub>3 and were incubated with and then digested by the on the left-hand side of e of the DNA sequence a fragment was labeled the bottom strand with w enzyme. Concentra- a was 4 · 10<sup>-6</sup> M (base of the peptide were as M; lane 2, 1.2 · 10<sup>-6</sup> M; 3 · 10<sup>-7</sup> M. Lane 0, con- e free DNA. Concen- d for digestion of naked es with the peptide were respectively. A+G, posi- equence cut by formic operator sites O<sub>R</sub>1, O<sub>R</sub>2 dooperator sites serving eptide II are marked on e gel.

r complexes of dis- ed DNA oligomers ers partially com- ide. As can be seen molar peptide/DNA bout 1, the fraction ble for binding of ual to about 60%, in e estimate obtained cal mechanical on- dicates that the pro- self-consistent values eters.

Binding of Peptide II

her peptide II is cap- g any particular se- l DNA fragments of e sequence footprint- ere performed using ecreatic DNase I or lated from walleye Figure 10 shows the nt pattern for the bot-

tom strand of the 306 bp DNA fragment containing specific interaction sites for the bacteriophage 434 repressor and *cro* proteins. Several regions of protection against cleavage by acidic DNase are observed on the DNA at peptide concentration of about 1.2 · 10<sup>-6</sup> M (lane 2). Three of them coincide with the operator sites O<sub>R</sub>1, O<sub>R</sub>2 and O<sub>R</sub>3. Cleavage protection is also detected between nucleotides 45-66, 141-154, 162-173, 192-204 and 210-228.

The boundaries of some protected regions are not well-defined due to a low frequency of enzyme cutting in these regions. Inspection of the gel patterns presented in Figure 10 shows that at low peptide concentrations isolated regions of protection are large enough to serve as interaction sites for the peptide (a bound peptide molecule appears to occupy at least 14 base pairs). The broad protected zone in positions 45-66 seems in fact to contain two closely-spaced sites of blockage. Comparison of the nucleotide sequences in the protected DNA regions with the sequence of the O<sub>R</sub>3 operator shows that all the protected regions contain four base pair sequences similar to those present on the left and right halves of the O<sub>R</sub>3 operator. The consensus sequence for twelve 434 operator halfsites is known to contain a critical tetranucleotide sequence 5'-A C A A-3' (10,11). This sequence motif with one mismatch is present in positions 142-145, 162-165, 193-196 and 209-212 on the top strand (Figure 11). It is also present with one mismatch in positions 42-52 and 219-222 on the bottom strand. In the right half of O<sub>R</sub>3 site, the adenine in position 4 of the consensus sequence is replaced by a guanine. The same sequence with one mismatch is present in positions 46-49, 156-159, 203-206 and 219-222 on the bottom strand. Interestingly, protected regions contain pairs of the characteristic four base pair sequences related by two-fold symmetry and separated by six base pairs. Examples are the pseudooperator sites O<sub>P</sub>1, O<sub>P</sub>2 and O<sub>P</sub>3 located in the 434 *cro* gene (Figure 11).

Figure 12 shows typical DNase I footprinting patterns generated by peptide II on the

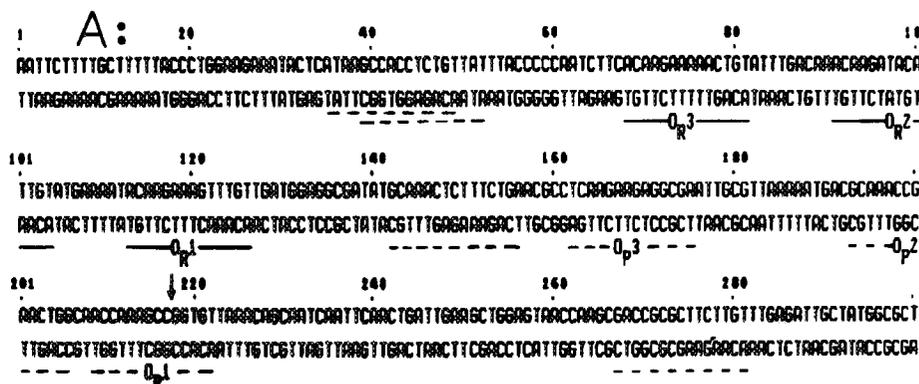


Figure 11: Sequence of the DNA fragment A. Operator sites O<sub>R</sub>1, O<sub>R</sub>2 and O<sub>R</sub>3 recognized by the peptide II on the fragment A are underlined by solid lines. Other DNA regions serving as preferential interaction sites for the peptide II on the DNA are underlined by dashed lines. Fragment B contains a part of the nucleotide sequence of the restriction fragment A (nucleotides 1 to 217). Vertical arrow shows the 3'-end of the upper strand of the fragment B.

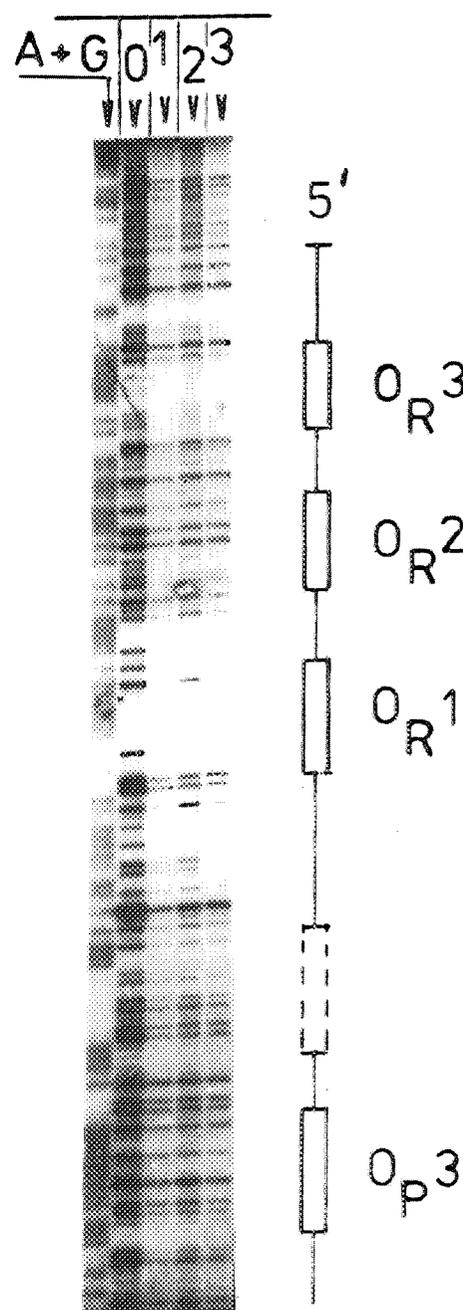


Figure 12: Autoradiograph of the DNase I protection pattern of the upper strand of DNA fragment B in the presence of peptide II. The DNA fragment was labeled selectively at the 3'-end of the upper strand with [ $\alpha$ - $^{32}$ P] CTP and Klenow fragment. Concentration of carrier DNA was  $3 \cdot 10^{-5}$  M (base pairs). Concentration of the peptide were as follows: lane 0, none; lane 1,  $8.4 \cdot 10^{-6}$  M; lane 2,  $4.2 \cdot 10^{-6}$  M; lane 3,  $2.1 \cdot 10^{-6}$  M. On the right-hand side of the gel the operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  and some other interaction sites for the peptide are indicated.

top strand of a DNA fragment B. Like fragment A, this DNA fragment contains operators  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$ . However, it does not contain the right half of pseudooperator  $O_{P1}$  (Figure 11). The peptide-induced patterns of protection observed for two different enzymes, DNase I and acidic DNase, correlate well, in spite of difference in experimental conditions. We interpret the peptide-induced protection of operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  (Figure 12) as evidence for binding of the peptide at or near the protected regions. Evidently, the peptide molecules occupy operators  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  and reduce accessibility of phosphodiester bonds in the operators and adjacent DNA regions to cleavage by DNase I. In the presence of  $5 \cdot 10^{-6}$  M peptide clear protection is observed in the DNA regions between nucleotides 30-60 and 140-156. It should be noted that in addition to these well-defined protection zones a gradual reduction in cleavage rate is observed in other DNA parts upon increasing the peptide concentration. Evidently, non-specific binding of the peptide to different nucleotide sequences can also inhibit enzyme activity. In Figure 13 representative densitometric scans of the autoradiogram shown in Figure 10 are presented. From a comparison of cleavage patterns obtained in the presence and absence of peptide II regions of DNA protected by the peptide can

graph of the DNase I pro-  
 pper strand of DNA frag-  
 of peptide II. The DNA  
 electively at the 3'-end of  
 $\alpha$ -<sup>32</sup>P] CTP and Klenow  
 on of carrier DNA was  
 Concentration of the pep-  
 e, none; lane 1,  $8.4 \cdot 10^{-6}$   
 lane 3,  $2.1 \cdot 10^{-6}$  M. On the  
 el the operator sites  $O_{R1}$ ,  
 other interaction sites for  
 ed.

JA fragment B. Like  
 DNA fragment con-  
 $O_{R2}$  and  $O_{R3}$ . How-  
 contain the right half  
 $O_{P1}$  (Figure 11). The  
 attens of protection  
 different enzymes,  
 ic DNase, correlate  
 erence in experimen-  
 nterpret the peptide-  
 n of operator sites  
 3 (Figure 12) as evi-  
 of the peptide at or  
 i regions. Evidently,  
 iles occupy operators  
 and reduce accessi-  
 diester bonds in the  
 acent DNA regions  
 ase I. In the presence  
 de clear protection is  
 NA regions between  
 nd 140-156. It should  
 ddition to these well-  
 1 zones a gradual re-  
 ge rate is observed in  
 upon increasing the  
 tion. Evidently, non-  
 of the peptide to dif-  
 sequences can also  
 ctivity. In Figure 13  
 nsitometric scans of  
 n shown in Figure 10  
 om a comparison of  
 obtained in the pre-  
 e of peptide II regions  
 d by the peptide can

be easily assigned. Difference scans can be generated by subtracting the naked DNA cutting pattern from that of the peptide-DNA complex (not shown). The height of each peak in difference scans may serve as a measure of the extent of protection at a given position on DNA. To facilitate comparisons of the relative affinities shown by peptide II for operators  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  and pseudooperator  $O_{P1}$  densitometric scans of gel patterns are obtained in the presence of various peptide concentrations. In the presence of  $6 \cdot 10^{-7}$  M peptide a prominent protection zone is centered on the pseudooperator  $O_{P1}$  (Figures 10 and 13). A weak blockage of bonds in  $O_{P1}$  is detectable even in the presence of  $3 \cdot 10^{-7}$  M peptide. By contrast, no protection is observed in  $O_{R3}$  under the same conditions. At concentration of  $6 \cdot 10^{-7}$  M cleavage rate is not changed in positions 67, 70, 78 and 79. It slightly increases in positions 64, 65 and 71. However, reduction in cleavage of phosphodiester bonds in  $O_{R3}$  is observed at peptide concentration as high as  $1.2 \cdot 10^{-6}$  M. We interpret these observations as indicating that peptide II binds more strongly to  $O_{P1}$  than to  $O_{R3}$ . To estimate the affinity of the peptide to  $O_{P1}$  the extent of protection in position 222 on the DNA was measured as a function of the peptide concentration (data are not shown). Interpolation was used to determine the concentration at which 50% protection against cleavage took place. Binding constant of the peptide to the  $O_{P1}$  was estimated to be approximately  $1 \cdot 10^6 \text{ m}^{-1}$ . It is about one order of magnitude greater than the affinity shown by the peptide for a synthetic operator  $O_I$ .

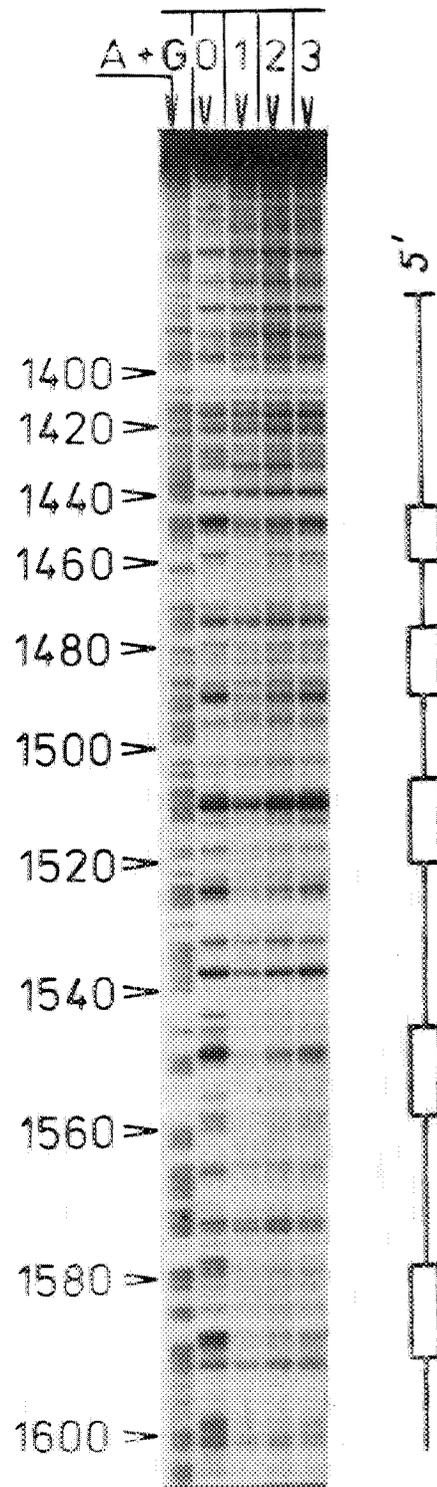
Inspection of densitometric scans also shows that at peptide concentration of  $6 \cdot 10^{-7}$  M reduction in cleavage of phosphodiester bonds takes place in  $O_{R2}$  (positions 101, 104 and 106) and  $O_{R1}$  (positions 112, 113, 122 and 123). The height of an intense peak arising from cleavage of DNA at position 126 (not shown) also decreased in the presence of peptide. We note that nucleotides occupying identical positions in operators  $O_{R1}$  and  $O_{R2}$  are protected in the presence of  $6 \cdot 10^{-7}$  M peptide, whereas protection is not observed in  $O_{R3}$  under the same conditions. This means that affinities shown by the peptide for binding to the  $O_{R1}$  and  $O_{R2}$  are greater than that shown for binding to the  $O_{R3}$  site. To compare relative affinities of the peptide for binding to  $O_{R1}$ ,  $O_{R2}$  and  $O_{P1}$  the extents of protection of adenine residues in positions 104, 126 and 222 were measured as functions of the peptide concentration. These residues occupy identical positions relative to a dyadic axis in each site. From Figure 13 it can be found that ratio of the absorbance reading at position 222 in the presence of peptide (A) to that measured in the absence of peptide ( $A_0$ ) is equal to 0.68 at the peptide concentration of  $6 \cdot 10^{-7}$  M. For cleavage at position 104 ( $O_{R2}$ ) the quantity  $A/A_0$  is equal to 0.83 under the same conditions. Digestions at positions 104 and 126 were equally sensitive to the presence of the peptide. This analysis shows that at low extents of binding the peptide binds more strongly to the pseudooperator  $O_{P1}$  than to the operators  $O_{R1}$  and  $O_{R2}$ . It is important to realize that 14 bp sequences homologous to the 434 operator sites are not the only sequences recognized by the nonlinear peptide on DNA. Footprinting analysis has identified a strong binding site for the peptide between nucleotides 141-156. The observed region of protection includes a 15 bp sequence in which the outer tetranucleotides are homologous to those present on the left and right halves of the  $O_{R3}$  operator. However, these sequences are related by an approximate dyad axis passing through a base pair at position 149 and are separated by seven base pairs. Evidently, flexibility of





**Figure 13:** Densitometric scans of the gel regions corresponding to the operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  and pseudooperator  $O_{p1}$  in the presence and absence of peptide II. Densitometric scans of the autoradiogram shown in Figure 10 are presented. Peak positions are given relative to the position of the uncleaved DNA. The ordinates represent absorbance readings. The height of each peak is a measure of the extent of cleavage at a particular position on the DNA. C is the peptide concentration.

the linker chains between the DNA-binding domains in the peptide molecule allows it to recognize pseudosymmetrical DNA sequences, irrespective of whether the critical tetranucleotides are separated by 6 or 7 base pairs. It should be noted that rate of cleavage increased at certain positions on the DNA in the presence of the peptide. At peptide concentration of  $6 \cdot 10^{-7}$  M enhanced cutting is observed in positions 64, 65, 71, 115 and 116. However, this effect disappears with further increasing the peptide concentration (Figure 13). At peptide concentration as high as  $1.2 \cdot 10^{-6}$  M reduction in cleavage is observed in all these positions, whereas enhanced cutting is seen in positions 170, 180, 211 and 230. Enhanced cutting in the presence of peptide may reflect a peptide-induced DNA structural change or can simply arise from alterations in the relative concentrations of free DNA and enzyme (42). At low extents of peptide binding the amount of enzyme bound to DNA is not practically changed relative to its level in the peptide-free system, thus indicating that peptide binding causes a shift of enzyme to free DNA regions. Under these conditions, as shown by Ward *et al.* (42), fractional enhancements of cleavage must be nearly of the same magnitude for all unblocked sites on the DNA. This is not the case for cleavage patterns generated in the presence of peptide II. Evidently, enzyme redistribution mechanism is only partially responsible for the observed enhancement of cleavage. It alone is unable to explain variation in the rate of cleavage observed for all unblocked sites on the DNA. This suggests that effects associated with a peptide-induced DNA structural change play an important role. CD measurements indicate that there is no gross alteration in the DNA structure upon peptide binding. However, these results do not rule out the possibility that peptide may induce a bend in the DNA upon binding. Previous studies on binding of small  $\beta$  structure-forming peptides to DNA have shown that these peptides can induce formation of DNA loops at low extents of binding and cause side-by-side association of DNA fragments at higher occupancies (43,44). It is also known that one *lac* repressor tetramer can bind to the two *lac* operators on one DNA fragment, thus causing the intervening DNA to form a stable loop (45). Formation of a loop is known to be favored by correct phasing of two *lac* operators and low concentrations of both components of the reaction. In the presence of excess DNA over protein two DNA fragments can bind to one *lac* repressor tetramer (45). We suggest that peptide II, like *lac* repressor, may interact with one operator site via two DNA-binding domains, and may contact a pseudooperator site on the same DNA fragment with the remaining two DNA-binding domains. The peptide may also bind to two DNA fragments simultaneously forming structure of sandwich-type. Preliminary electron microscopy studies shows that peptide binding can induce compaction of DNA. To demonstrate that protection patterns observed in the presence of peptide II are unique and form only on the DNA fragment containing 434 *cro* gene and operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$ , footprinting experiments were performed with a 410 bp nonspecific DNA fragment from pUC9 plasmid. Figure 14 shows the peptide II footprint generated on the upper strand of the DNA fragment. Clear protection is observed at certain DNA regions,



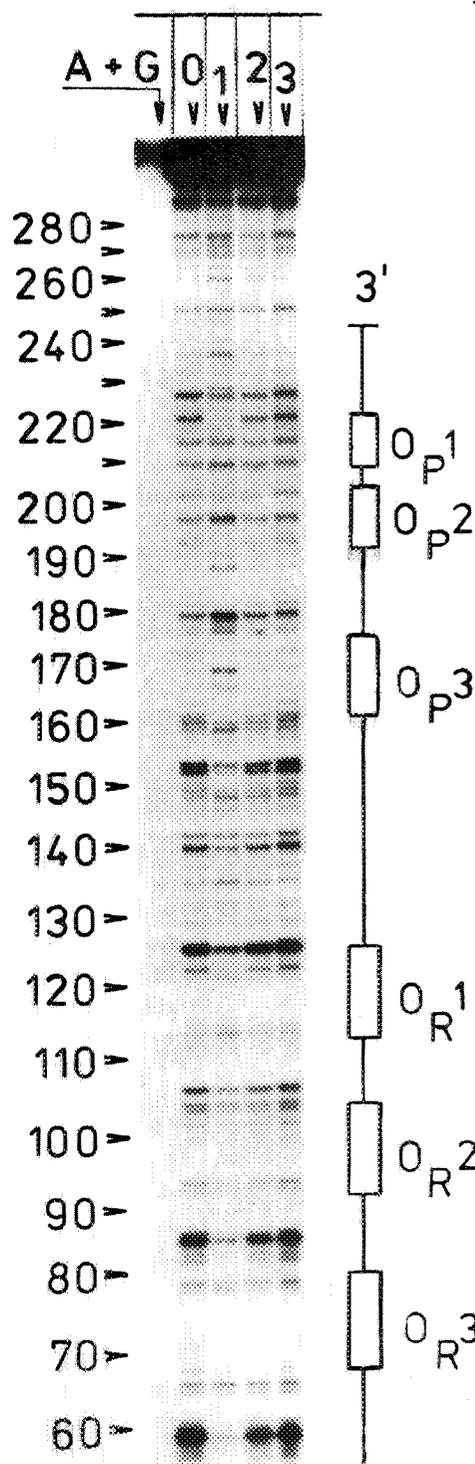
**Figure 14:** Digestion of the DNA fragment C by acidic DNase in the absence and presence of peptide II. The numbers on the left-hand side of the gel correspond to those of the DNA sequence of fragment C (Figure 15). The DNA fragment was labeled selectively at the 3'-end of the upper strand with [ $\alpha$ - $^{32}$ P] dGTP and Klenow enzyme. The  $^{32}$ -P-end-labeled fragment C and unlabeled pUC9 DNA were incubated for 20 min at 20°C and digested for 60 seconds by the acidic DNase. Concentration of pUC9 DNA was  $8.5 \cdot 10^{-6}$  M (base pairs). Concentrations of the peptide were: lane 0, none; lane 1,  $6 \cdot 10^{-6}$  M; lane 2,  $2 \cdot 10^{-6}$  M; lane 3,  $7 \cdot 10^{-7}$  M. On the right-hand side of the gel some isolated interaction sites for peptide II on the restriction fragment are indicated.

indicating that peptide II binds at preferred sites on the DNA. However, structural details of the protection pattern differ significantly from those observed above for complexes with the DNA fragment containing operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$ . Inspection of nucleotide sequences in isolated binding sites for the peptide on DNA shows that they all exhibit sequence similarities with the 434 operator sites. One example is the sequence 5'-TCAAGAAGATCC TTT-3' (positions 1446-1459). In this sequence each of the two outer tetranucleotides has only one mismatch with the corresponding sequences in  $O_{R3}$ . Clear protection of phosphodiester bonds at this site is observed in the presence of  $2 \cdot 10^{-6}$  M peptide. Other isolated protected regions on the DNA fragment also contain sequences which show similarities with the consensus sequence for phage 434 operators (see Figure 15).

#### *Sequence-Specific Binding of Peptide III to DNA*

The ability of peptide III to bind to the operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  is demonstrated in Figure 16. The footprinting patterns generated by peptides II and III on the same DNA may





**Figure 16:** Digestion of the DNA fragment A by acidic DNase in the presence and absence of peptide III. Concentration of carrier DNA in the reaction mixture was  $1.6 \cdot 10^{-5}$  M (base pairs). Concentrations of the peptide were as follows: lane 0, none; lane 1,  $2.4 \cdot 10^{-3}$  M; lane 2,  $7.9 \cdot 10^{-6}$  M; lane 3,  $2.7 \cdot 10^{-6}$  M. To obtain similar extents of digestion of naked DNA and peptide-DNA complexes the enzyme concentrations in the reaction mixtures containing peptide-DNA complexes were 2 to 4 times greater than that used for digestion of naked DNA ( $0.5 \mu\text{g}/\text{ml}$ ). Experimental conditions were identical to those in Figure 10.

model of B form DNA, the enhanced cutting sites all lie on one face of the DNA with a periodic repeat of about 10 base pairs. These features are typical of footprints generated for nonspecific peptide-DNA complexes. The peptide appears to bind to one side of the DNA helix in an ordered and, most likely, cooperative manner.

#### Discussion

The results reported in the present work show that peptide II binds to DNA and recognizes operator sites O<sub>R1</sub>, O<sub>R2</sub> and O<sub>R3</sub>, with which 434 repressor and *cro* proteins interact selectively. Unlike the 434 *cro* protein, the peptide II possesses a greater affinity to a pseudooperator O<sub>P1</sub>, than to the operator O<sub>R3</sub>. In addition, it binds tightly to a few sites, located in the *cro* gene, that, to our knowledge, have not yet been identified as strong interaction sites for the 434 repressor and *cro* proteins. Analysis of the footprint patterns generated by the peptide on the two DNA fragments shows that peptide II recognizes pseudosymmetrical sequences 5'-ACA(W)<sub>n</sub>YTG T-3', where W is an arbitrary nucleotide; Y is a pyrimidine; n is equal to 6 or 7. It should be noted that pseudooperator O<sub>P1</sub> contains a sequence motif 5'-ACCA-3' on its left-handed side, in-

the DNA fragment A by the presence and absence of carrier DNA in the presence of  $10^{-5}$  M (base pairs). The concentrations of peptide were as follows: lane 1,  $10^{-5}$  M; lane 2,  $7.9 \cdot 10^{-6}$  M. To obtain similar extents of binding, the concentrations of DNA and peptide-DNA complexes in the reaction were higher than those used for the control (1.5  $\mu$ g/ml). Experimental conditions are the same as those in Figure 1.

peptide II, the enhanced binding to one face of the DNA repeat of about 10 base pairs are typical of nonspecific peptides. The peptide approaches the DNA from one side and, most likely,

in the present work peptide II binds to DNA and recognizes operator sites  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  of the repressor and *cro* protein. Unlike the other peptides, peptide II possesses a pseudooperator site  $O_{R3}$ . In addition to a few sites, located near our knowledge, identified as strong binding sites for the 434 repressor protein. Analysis of the footprints of the peptide-DNA complexes shows that peptide II binds to pseudosymmetrical sites  $C A A (W)_n Y T G$  where W is an arbitrary nucleotide; n is equal to 6 or 7. The pseudooperator sequence motif 5'-A

on the left-handed side, in-

stead of the sequence 5'-A C A A -3' recognized by the 434 repressor and *cro* proteins. Since peptide II can recognize operator-like sequences on DNA, we suggest that a pair of its bihelical motifs recognize and bind to the right and left halves of a pseudosymmetrical operator, in spite of chemical modifications introduced in the bihelical motifs. Evidently, flexible linker chains which connect the DNA-binding domains with the C-terminal crosslinker, allow for a pair of the bihelical motifs to contact two patches of DNA, one located at each end of the operator. Our observations show that a linear peptide containing the modified helix-turn-helix motif of the 434 *cro* interacts selectively with the operator-like sequences on DNA and, presumably, recognizes the operator halfsite. This is consistent with the idea that  $\alpha$  helix-turn- $\alpha$  helix motif of a bacterial repressor is implicated in specific interactions with DNA (1,46).

The affinities, with which peptides II and III bind to the operators  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$ , suggest that a single peptide II molecule contacts the two operator halfsites. Consistent with this are our observations that peptide II protects from nuclease digestion a longer DNA region, as compared with that protected by peptide III. Considerable flexibility inherent in the peptide II appears to be responsible for its capacity to recognize pseudosymmetrical sequences, in which critical tetranucleotides are separated by 6 or 7 base pairs. Synthesis of peptide II analogs with improved binding specificity requires replacement of flexible linker chains in the peptide molecule by a more rigid linker. Design and synthesis of sequence-specific DNA-binding non-linear peptides offer new possibilities for *de novo* design of peptides and proteins with repressor-like properties. It can be easily seen that with maintenance of the connectivity in peptide II replacement of helix-turn-helix motifs of the 434 *cro* by bihelical motifs of another bacterial repressor enables one to construct a non-linear peptide that may recognize different nucleotide sequences on DNA. Synthesis can be simplified, if the four-arm crosslinker is replaced by two-armed one, with each arm bearing a single helix-turn-helix motif of a bacterial repressor. It is also possible to link helix-turn-helix motifs of two different repressors. These constructs might possess a composite sequence-specificity, distinct from the specificity shown by each individual protein.

In aqueous solutions peptides I, II and III exist predominantly in disordered conformations (8,9). CD spectral changes observed for the two peptides in the presence of 20% TFE reflect a partial formation of  $\alpha$ -helical structure. This indicates that a relatively small change in solvent conditions is required to induce transition from a disordered to the  $\alpha$ -helical conformation. CD studies show that peptides II and III undergo conformational changes upon binding to DNA. Our present and previous observations (5,8,9) indicate that CD spectrum of a peptide-DNA mixture is not simply a sum of the CD spectrum of the free peptide and spectrum of the corresponding nucleic acid. The difference CD pattern, obtained by subtracting the CD spectrum of nucleic acid from the spectrum of peptide-DNA mixture is reminiscent of a  $\beta$ -sheet CD pattern. The difference spectra generated for peptide II in the presence of calf thymus DNA, poly(dA)·poly(dT), poly[d(AC)]·poly[d(GT)] and operator DNA fragments are all very similar, indicating that bound peptide molecules take up similar conformations in all these complexes. The difference CD spectra for

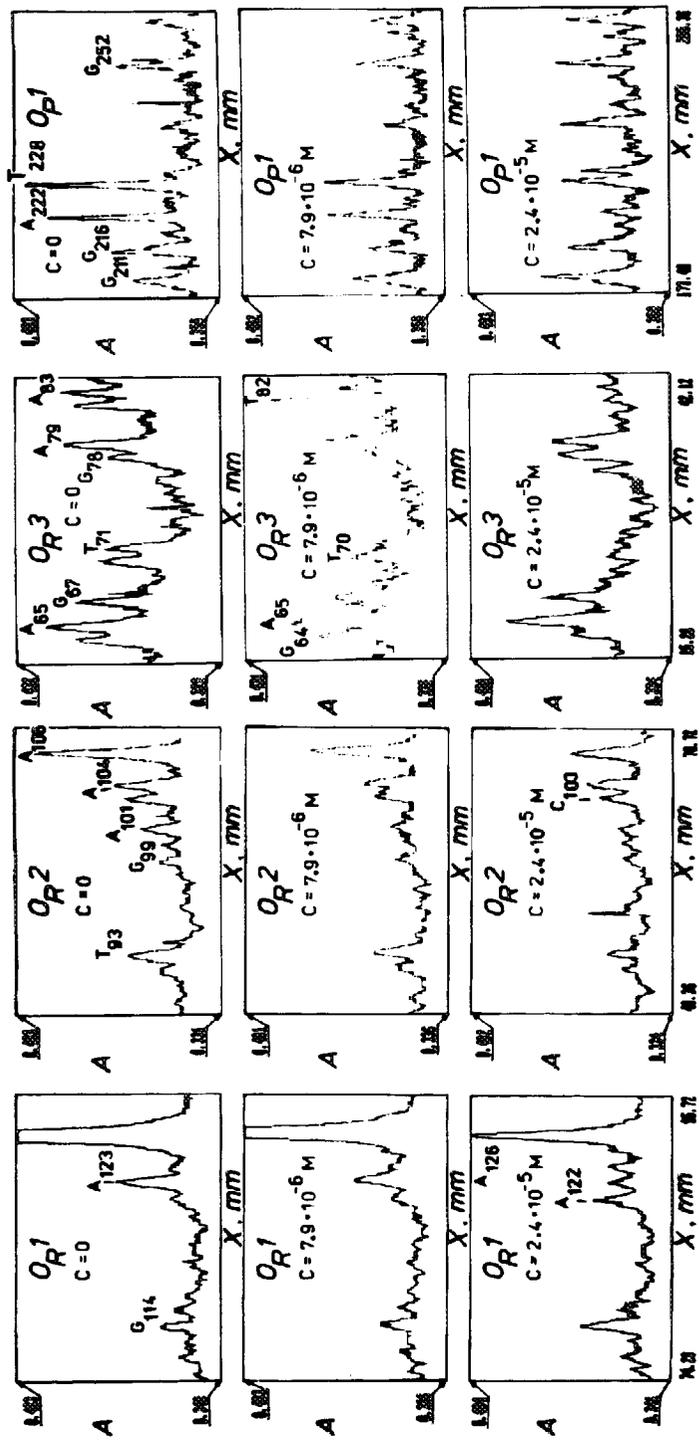




Figure 17: Densitometric tracing of the gel patterns shown in Figure 16. Densities at particular positions of the operators  $O_{R1}$ ,  $O_{R2}$  and  $O_{R3}$  and pseudoooperator  $O_{p1}$  are measured as functions of the peptide concentration (C).

peptides I and III also resemble a  $\beta$ -like CD pattern. We interpret these observations to mean that  $\alpha$ - $\beta$  conformational transition takes place in the bihelical motifs of peptides I, II and III upon binding to DNA. Evidently, each helix-turn-helix sub-structure undergoes conversion into a motif  $\beta$  strand-turn- $\beta$  strand upon complex formation with DNA.

It should be noted that  $\alpha$ - $\beta$  transition is a widespread phenomenon. Some peptides may exist in  $\alpha$ - or  $\beta$ -conformation, depending on the degree of polymerization, type of solvent, temperature and other conditions (32-34,47-50). The free energy difference between the  $\alpha$ -helix and  $\beta$ -sheet conformations of poly-L-lysine is very small (50). The induction of  $\alpha$ - $\beta$  transition appears to depend on the peptide length and amino acid sequence (47,48). Votavova *et al.* have found that polymers  $(\text{Lys Ala Ala})_n$  and  $(\text{Lys Leu Ala})_n$  ( $n=10$ ) assume  $\alpha$ -helical structures in aqueous solution in the presence of 30% TFE and that the helical structure is preserved upon binding of the two peptides to DNA (51). In contrast with the results obtained for the complexes of peptides I, II and III with DNA, difference CD spectra for  $(\text{Lys Ala Ala})_n$  and  $(\text{Lys Leu Ala})_n$  resemble the CD spectrum for  $\alpha$ -helical structure. Only a small amount (if any) of  $\beta$ -structure is induced upon binding of these peptides to DNA, presumably, reflecting a low  $\beta$  structure-forming potential of these amino acid sequences.

The induction of helix- $\beta$  transition upon binding of peptides I, II and III to DNA implies that  $\beta$ -conformation is a more favored structure than  $\alpha$ -helix for interaction of these peptides with DNA. The helix-turn-helix motif of a bacterial repressor has a structural feature that may favor helix- $\beta$  transition. The backbone configuration of eight consecutive residues in the helix-turn-helix motif, including the turn between helices  $\alpha_2$  and  $\alpha_3$  can be described as  $\dots\alpha_R\alpha_R\alpha_R\gamma_R\alpha_L\beta\beta\alpha_R\dots$ , where  $\alpha_R$ ,  $\gamma_R$ ,  $\alpha_L$  and  $\beta$  denote residues with main chain dihedral angles  $\phi$  and  $\psi$ , corresponding to the right-handed  $\alpha$ -helix ( $\alpha_R$ ),  $3_{10}$  helix ( $\gamma_R$ ), left-handed  $\alpha$ -helix ( $\alpha_L$ ) and  $\beta$ -stranded ( $\beta$ ) regions on Ramachandran map (52,53). Since the helix-turn-helix motif contains a stable "nucleus" of  $\beta$ -stranded structure near the N-terminus of helix  $\alpha_3$ , we believe that helix-turn-helix motif is a more favorable structure for cooperative helix- $\beta$  transition than a long uninterrupted  $\alpha$  helix with the same number of residues as in helices  $\alpha_2$  and  $\alpha_3$ . The helix- $\beta$  conformational transition in the bihelical motif may lead to the formation of a  $\beta$ -hairpin. The backbone configuration of a standard  $\beta$ -hairpin with a four residue loop exhibits a structural similarity with the above described region in the helix-turn-helix motif. The characteristic configuration of the  $\beta$ -hairpin has been shown to have the following sequence of  $\phi$ - $\psi$  value types for eight consecutive residues, including the loop between the two  $\beta$  strands:  $\dots\beta\alpha_R\alpha_R\gamma_R\alpha_L\beta\beta\dots$  (53). The backbone configuration of the helix-turn-helix motif of phage  $\lambda$  *cro* protein (residues 20-26) and that of a  $\beta$ -hairpin in bovine trypsin inhibitor (residues 24-30) can be superimposed with a root mean square deviation of about 0.6 Å (54), indicating that 7 out of 21  $C^\alpha$  atoms in the helix-turn-helix motif occupy spatial positions appropriate for formation of a  $\beta$ -hairpin. The backbone atoms of the bihelical motifs of 434 *cro* (residues 23-29) and *trp* repressor (residues 72-78) can be superimposed

on those of residues 24-30 of the  $\beta$ -hairpin with r.m.s. differences of about 0.6 Å (54). This structural similarity suggests that the helix-turn-helix motif may have a high propensity to helix- $\beta$  transition and transformation into a  $\beta$ -hairpin. The  $\alpha$ -helical structure is known to be stabilized by side chain-side chain interactions, including van der Waals contacts, hydrogen bonding and ion pair formation between residues in positions  $i$ ,  $i+3$  and  $i+4$  on  $\alpha$  helix (55,56). Interaction between protein and DNA may greatly perturb intrahelical side chain-side chain interactions and may induce destabilization of helices  $\alpha_2$  and  $\alpha_3$  and their conversion into  $\beta$ -strands. A precedent for the helix-coil transition induced in the presence of DNA can be found in recent studies showing that the histone core has a lower  $\alpha$  helix content and higher random coil content upon binding to DNA (57). A decrease in  $\alpha$  helix content accompanied by an increase in  $\beta$  strand content has been detected in the gp32 protein of bacteriophage T<sub>4</sub> upon binding oligonucleotides (58).

Our observations that substantial conformation changes are induced in the helix-turn-helix motifs of peptides II and III upon binding DNA are unexpected in the light of recent crystallographic studies of repressor-operator complexes (10,11,59). In crystalline complexes, the backbone configurations of helix-turn-helix substructures are very similar to those found in the uncomplexed repressors (10,11,24,59-61). This indicates that modes of binding of the bihelical motifs of peptides II and III to DNA differ from those observed in crystalline complexes of 434 repressor headpiece and 434 *cro* with synthetic operators (10,11). The binding of peptides II and III to DNA appears to be coupled with the  $\alpha$ - $\beta$  conversion in the two-helix motif. Since bound peptides contain virtually no  $\alpha$ -helical structure, alternative arrangements are modeled in which two antiparallel  $\beta$ -stranded chains are inserted either in the major or minor DNA groove. A precedent for the former arrangement is found in the crystalline complex of methionine repressor with a synthetic operator DNA fragment (62). It seems plausible that bihelical motifs of peptides II and III are transformed into  $\beta$ -hairpins that interact with the functional groups of DNA base pairs exposed in the major groove. An alternative model is based on our previous observations that  $\beta$  structure-forming peptides can bind in the minor DNA groove in a sequence-specific manner (5-8). These peptides exist in disordered conformations in water solution, but take up  $\beta$ -like conformations upon complex formation with DNA (5-8). The DNA undergoes subtle changes in its helical parameters upon binding of peptide. Our present data show that peptides II and III displace the minor groove-binding antibiotic distamycin A from poly(dA) · poly(dT) and synthetic DNA oligomers O<sub>I</sub> and O<sub>II</sub>. We propose, as a working model, that binding of peptides II and III to DNA is accompanied by the helix- $\beta$  transition in the two-helix motifs and insertion of corresponding  $\beta$ -hairpins in the minor groove. Although other interpretations are possible and detailed structural information is not yet available, molecular model-building studies and calculation of energy-minimum conformations show that backbone NH and CO groups of the two antiparallel  $\beta$ -stranded chains can serve as hydrogen bond donors and acceptors for interaction with the functional groups of DNA base pairs exposed in the minor groove (38,63). This recognition mechanism is observed in the crystalline complex of quinoxaline antibiotic triostin A with a self-complementary hexanuclotide 5' -C G A T C G-3' (64). We believe that this mechanism is not restricted solely to DNA complexes with

about 0.6 Å (54).  
 may have a high  
 in. The  $\alpha$ -helical  
 tions, including  
 between residues  
 rotein and DNA  
 and may induce  
 nds. A precedent  
 e found in recent  
 d higher random  
 nt accompanied  
 ? protein of bac-

ced in the helix-  
 xpected in the  
 olexes (10,11,59).  
 -helix substruc-  
 (10,11,24,59-61).  
 ides II and III to  
 repressor head-  
 eptides II and III  
 elix motif. Since  
 ve arrangements  
 ted either in the  
 ment is found in  
 o operator DNA  
 es II and III are  
 ps of DNA base  
 on our previous  
 or DNA groove  
 ordered conform-  
 mplex formation  
 parameters upon  
 III displace the  
 oly(dT) and syn-  
 , that binding of  
 in the two-helix  
 oove. Although  
 ation is not yet  
 nergy-minimum  
 o antiparallel  $\beta$ -  
 s for interaction  
 or groove (38,63).  
 x of quinoxaline  
 C G A T C G-3'  
 complexes with

the antibiotics of triostin family. In this regard it should be noted that there is a large group of proteins that use  $\beta$ -stranded arms to bind to DNA (65,66). Among these proteins the *E.coli* integration host factor appears to be an example of the protein that recognizes specific base pair sequences in the minor DNA groove (67).

The binding specificity shown by certain  $\beta$  structure-forming peptides can be described in terms of the recognition code formulated on the basis of the above-described model for interaction of the two antiparallel  $\beta$ -stranded chains with DNA (6,8). The proposed code implies that side chain-backbone interactions within a deformed  $\beta$ -sheet control both the strength and specificity of hydrogen bonding interactions between the backbone NH and CO groups of the two antiparallel  $\beta$  strands and DNA base pairs (38,40,63).

According to the proposed code, residues Ser, Thr, Asn, Gln, His and Cys (potential hydrogen bond donors) can serve as AT-coding residues. Their spatial positions in the specific complex correspond to the array of AT-base pairs. Amino acid residues with inert side chains (Gly, Ala, Val, Leu, Ile, Phe, Met) and some other residues code for GC-pairs. The adjacent coding residues occupy positions  $i$  and  $i \pm 2$  along peptide chain. In the specific complex, the direction of peptide backbone from the N-terminus to the C-terminus in each  $\beta$ -strand coincides with the C3'-C5' direction of the adjacent polynucleotide strand, as found in the crystalline complex between triostin A and the hexanucleotide 5'-C G A T C G-3' (64). These rules are supported by the following experimental data. The tripeptide H-Val-Val-Val-NH-NH-Dns (where Dns is a residue of 5-dimethylaminonaphthalene-1-sulfonic acid) in the  $\beta$ -associated form binds more strongly to poly(dG) · poly(dC), than to poly(dA) · poly(dT) and poly[d(GC)] · poly[d(GC)] (5,6). The aromatic chromophore linked to the C-terminus of trivaline has been shown not to affect the binding specificity of the peptide (6,68,69). A pair of the peptides H-Lys-Gly-Val-Cys-Val-NH-NH-Dns bridged by a disulphide bond exhibits 20 times greater affinity for binding to poly(dG) · poly(dC) than to poly(dA) · poly(dT) (70). Recently, a synthetic sequence-specific ligand has been constructed, in which two netropsin-like fragments are linked covalently to the C-termini of a pair of peptides H-Gly-Cys-Gly-Gly-Gly-Gly joined together by a disulfide bond. This bis-netropsin binds in the minor groove and possesses a composite binding specificity: its peptide chains in a  $\beta$ -associated form interact with  $3'-CC$  steps on DNA, whereas each netropsin-like fragment recognizes three consecutive AT-base pairs (69). Furthermore, using the above described code rules, a 20-residue peptide has been designed, in which the distribution pattern of threonine and valine residues along the peptide chain corresponds to the array of AT- and GC-pairs in the idealized halfsite of operator  $O_R$  of phage  $\lambda$  (8). This peptide binds in the minor groove and recognizes operator sites  $O_{R1}$  and  $O_{R2}$  (8). However, it does not bind to a modified operator  $O_{R3}$  in which four base pairs are changed.

The binding specificity shown by peptides II and III also agrees with the code proposed for description of specific interactions between  $\beta$ -structure-forming peptides and DNA. One can easily find that sequence Thr Gln Thr Glu Leu Ala Thr Lys, spanning residues 18 to 25 of the 434 *cro* protein, matches perfectly the critical tetranucleotide sequence  $5'-TTGT-3'$  recognized by the 434 repressor and *cro* proteins.

Here amino acid residues coding for AT and GC pairs are underlined by solid and dotted lines, respectively. In peptide III there is only one replacement (Glu21-Ala) in this sequence, which can affect the general affinity of helix-turn-helix motif for nucleic acids, but not so much binding specificity (71). In peptide II the two amino acid replacements (Thr20-Ala and Glu21-Lys) are introduced in this sequence. Threonine 20 is assumed to play a key role in recognition of TA pair in the second position of the tetranucleotide sequence. Since alanine serves as a GC-coding residue, one might expect that base pair sequence  $\begin{matrix} 5'-TGGT-3' \\ 3'-ACCA-5' \end{matrix}$  exhibits a better fit with the amino acid sequence in each arm of the peptide II than does the consensus sequence for the operator halfsite. The former sequence is present in the pseudo-operator  $O_{P1}$ , to which peptide II binds more strongly than to the operator  $O_{R3}$ . The same relationship, with appropriate modifications, is found for several other specific protein-DNA complexes, including the complex between *E. coli* tryptophan repressor and tryptophan operator (40). Generally, residues present either in helix 2 or helix 3 of a two-helix motif exhibit a perfect fit with a recognizable base pair sequence. We found that sequence Ala Thr Ile Thr Arg Gly Ser Asn Ser, containing residues 80-88 of *trp* repressor, matches perfectly the critical tetranucleotide sequence  $\begin{matrix} 5'-TAGT-3' \\ 3'-ATCA-5' \end{matrix}$  in the idealized tryptophan operator. As before, amino acid residues coding for AT- and GC-base pairs are underlined by solid and dotted lines, respectively. This correspondence implies that Thr 81 plays a key role in specific recognition of TA pair in the first position of the tetranucleotide sequence. Consistent with this are *in vivo* binding studies of mutant *trp* repressors (72). When applied to the cAMP-receptor protein, CAP, the same rules enable one to identify a sequence Thr Arg Gln Glu Ile Gly Gln Ile Val Gly (residues 168-177) that exhibits the best fit with the sequence  $\begin{matrix} 5'-ATGTG-3' \\ 3'-TACAC-5' \end{matrix}$  in the CAP interaction site (40).

In the crystalline *trp* repressor-operator complex no direct hydrogen bond and van der Waals contact is found between the amino acid residues of recognition helix and the functional groups of DNA base pairs in the major groove (59). Sigler *et al.* have proposed that the operator sequence is recognized indirectly through its effect on the geometry of the phosphate backbone. However, the observed structure may represent the geometry of a nonspecific repressor-operator complex (9). If the  $\alpha$  helix- $\beta$  transition is a prerequisite for formation of the specific repressor-operator complex, then sequence correlations described above for the *trp* repressor and 434 *cro* suggest that specificity determinants of the two proteins are similar, in spite of the observed difference in binding geometries of their helix-turn-helix motifs in crystals. It is possible that helix-turn-helix substructure is invented in the course of molecular evolution in order to decrease interaction surface between protein and DNA for initial nonspecific complexes. Helix- $\beta$  transition induced in two-helix motif upon complex formation between protein and DNA permits sliding of the relevant parts of the protein molecule along the DNA helix to search for a local minimum of binding energy (40). The search for a specific binding site on DNA can be represented as a series of association and dissociation processes coupled with  $\alpha$ - $\beta$  and  $\beta$ - $\alpha$  transitions in two-helix motifs. An important feature of the proposed model is that helices 2 and 3 of the two-helix motif are converted into  $\beta$  strands, each capable of interacting with DNA base pairs. This may explain our observations that removal of two residues from the N-terminus of helix  $\alpha_2$  (this is the case in four arms of peptide I leads to a

ined by solid and  
nent (Glu21-Ala)  
n-helix motif for  
II the two amino  
in this sequence.  
pair in the second  
as a GC-coding  
ts a better fit with  
es the consensus  
nt in the pseudo-  
operator O<sub>R</sub>3. The  
eral other specific  
tophan repressor  
helix 2 or helix 3  
air sequence. We  
ng residues 80-88  
uence 5'-TAGT-3'  
3'-ATCA-5' in  
s coding for AT-  
ctively. This cor-  
tion of TA pair in  
h this are *in vivo*  
e cAMP-receptor  
Arg Gln Glu Ile  
ith the sequence

en bond and van  
gnition helix and  
Sigler *et al.* have  
ugh its effect on  
ructure may rep-  
) . If the  $\alpha$  helix- $\beta$   
operator complex,  
d 434 *cro* suggest  
e of the observed  
in crystals. It is  
se of molecular  
d DNA for initial  
if upon complex  
arts of the protein  
ding energy (40).  
ed as a series of  
ntitions in two-  
elices 2 and 3 of  
interacting with  
of two residues  
ptide I leads to a

loss of binding specificity of the synthetic peptide (9). There is much interest in systematic study of various factors that may favor helix- $\beta$  conversion in a bihelical motif of a bacterial repressor. The fact that helices 2 and 3 are preserved in crystalline repressor-operator complexes can be explained as being due to either unfavorable conditions in crystals or sterical hindrance for helix- $\beta$  transition caused by contacts that residues in helices 2 and 3 make with the rest of the protein molecule. The crystals of *trp* repressor-operator complex were grown in the presence of 35% pentadiol, 50 mM NaCl and about 10 mM CaCl<sub>2</sub>. Crystallization buffer for 434 repressor headpiece-operator complex contained approximately 12% polyethylene glycol, 120 mM MgCl<sub>2</sub>, 2 mM spermine and 100 mM NaCl. These conditions differ significantly from those employed in our studies and might be unfavorable for helix- $\beta$  conversion.

Recently it has been found that difference CD spectrum for  $\lambda$  phage *cro* repressor complexed with a synthetic operator DNA fragment differs from the CD spectral profile of the uncomplexed protein at wavelength region of 190-240 nm (41). However, the observed CD pattern is unlike the difference CD spectrum for peptide II bound to the DNA oligomer O<sub>I</sub>. This suggests that considerable structural rearrangements are introduced in the  $\lambda$  *cro* protein upon binding DNA, although their nature remains to be unknown.

The helix-turn-helix motif seems to be an autonomous folding domain stabilized basically by interactions between residues in helices 2 and 3. In the bihelical motif of 434 *cro* glutamate 37 is hydrogen bonded to glutamine 19. In addition, Leu22 and Ile31 are separated by a short distance (24). If a helix- $\beta$  conversion takes place in the bihelical motif, the hydrophobic side chains of these residues may interact with each other and stabilize structure of  $\beta$ -hairpin (9). Our present results coupled with literature data suggest that residues in the two-helix motif may exist either in  $\alpha$  helix- or  $\beta$ - conformation when bound to DNA. Peptides II and III (or their analogs) appear to be good model systems for finding agents and environment conditions that can differentially stabilize  $\alpha$ - or  $\beta$ - conformations in specific peptide-DNA complexes.

#### Acknowledgments

We would like to thank Dr. V. Kozhemyako for a gift of preparation of acidic DNase. Synthesis of peptide I was performed by S.L. Grokhovsky during his work at Institute of Genetics Cologne University (FRG). We would like to thank Benno Muller-Hill for support, advices and valuable discussions.

#### References and Footnotes

1. Pabo, C.O. and Sauer, R.T., *Ann. Rev. Biochem.* 53, 293-321 (1984).
2. Bruist, M.F., Horvath, S.J., Hood, L.E., Steitz, T.A. and Simon, M.I., *Science* 235, 777-780 (1987).
3. Sluka, J.P., Horvath, S.J., Bruist, M.F., Simon, M.I. and Dervan, P.B., *Science* 238, 1129-1132 (1987).
4. Mack, D.P., Iverson, B.L. and Dervan, P.B., *J. Am. Chem. Soc.* 110, 7572-7574 (1988).
5. Streltsov, S.A., Khorlin, A.A., Surovaya, A.N., Gursky, G.V., Zasedatelev, A.S., Zhuze, A.L. and Gottikh, B.P., *Biofizika (USSR)* 25, 929-941 (1980).
6. Gursky, G.V., Zasedatelev, A.S., Zhuze, A.L., Khorlin, A.A., Grokhovsky, S.L., Streltsov, S.A., Surovaya, A.N., Nikitin, S.M., Krylov, A.S., Retchinsky, V.O., Mikhailov, M.V., Beabealashvilli, R.Sh. and Gottikh, B.P., *Cold Spring Harbor Symp. Quant. Biol.* 47, 367-378 (1983).

7. Vengerov, Y.Y., Semenov, T.E., Surovaya, A.N., Sidorova, N.Y., Streltsov, S.A., Khorlin, A.A., Zhuze, A.L. and Gursky, G.V., *J. Biomol. Struct. Dyn.* 6, 311-330 (1988).
8. Grokhovsky, S.L., Surovaya, A.N., Sidorova, N.Y., Votavova, H., Sponar, J., Fric, I. and Gursky, G.V., *Mol. Biol. (USSR)* 22, 1315-1333 (1988).
9. Grokhovsky S.L., Surovaya A.N., Sidorova N.Y. and Gursky G.V., *Molekul. Biol. (USSR)* 23, 1558-1580 (1989).
10. Wolberger, C., Dong, J., Ptashne, M. and Harrison, S.C., *Nature* 335, 785-795 (1988).
11. Aggarwal, A.K., Rodgers, D.W., Drottar, M., Ptashne, M. and Harrison, S.C., *Science* 242, 899-907 (1988).
12. Wharton, R.P., Brown, E.L. and Ptashne, M., *Cell* 38, 846-852 (1984).
13. Wharton, R.P. and Ptashne, M., *Nature* 316, 601-605 (1985).
14. Zimmer, Ch. and Wahnert, U., *Prog. Biophys. Mol. Biol.* 47, 31-112 (1986).
15. Barany, G. and Merrifield, R.B., *The Peptides* 6, 3-284 (1980).
16. Felix, A.M. and Jimenes, H.H., *Anal Biochem.* 52, 377-381 (1973).
17. Tam, J.P., Heath, W.F. and Merrifield, R.B., *J. Am. Chem. Soc.* 105, 6442-6455 (1983).
18. Ovchinnikov, Yu. A., Guryev, S.O., Krayev, A.S., Monastyrskaya, G.S., Skryabin, K.G., Sverdlov, E.D., Zhakaryev, V.M. and Bayev, A.A., *Gene* 6, 235-249 (1979).
19. Maniatis, T., Fritsch, E.F. and Sambrook, J., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory, Cold Spring Harbor, New York (1982).
20. Rasskazov, V.A. and Kozhemyako, V.B., *Dokl. Acad. Nauk USSR* 294, 1497-1500 (1987).
21. Maxam, A.M. and Gilbert, W., *Methods Enzymol.* 65, 449-560 (1980).
22. Chernov, B.K., Golova, Yu.B., Pomosgova, G.E. and Bayev, A.A., *Dokl. Akad. Nauk USSR* 291, 1131-1143 (1986).
23. Chang, C.T., Wu, C.M.C. and Yang, J.T., *Anal. Biochem.* 91, 13-31 (1978).
24. Mondragon, A., Wolberger, C. and Harrison, S.C., *J. Mol. Biol.* 205, 179-188 (1989).
25. Greenfield, N. and Fasman, G.D., *Biochemistry* 8, 4108-4116 (1969).
26. Brahms, S., Brahms, J., Spach, G. and Brack, A., *Proc. Nat. Acad. Sci. USA* 74, 3208-3212 (1977).
27. Goodman, E.M. and Kim, P.S., *Biochemistry* 28, 4343-4347 (1989).
28. Stokrova, S., Zimmerman, K., Blaha, K. and Sponar, J., *Coll. Czech. Chem. Commun.* 43, 2341-2350 (1978).
29. Woody, R.D., *The Peptides* 7, 15-114 (1988).
30. Baase, W.A. and Johnson, W.C., Jr., *Nucleic Acids Res.* 6, 797-814 (1979).
31. Alexeev, D.C., Lipanov, A.A. and Skuratovski, I.Ju., *Nature* 325, 821-823 (1987).
32. Sarkar, P.K. and Doty, P., *Proc. Nat. Acad. Sci. USA* 55, 981-989 (1966).
33. Satake, I. and Yang, J.T., *Biochem. Biophys. Res. Commun.* 54, 930-936 (1973).
34. Satake, I. and Yang, J.T., *Biopolymers* 15, 2263-2275 (1976).
35. Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E., *Proc. Natl. Acad. Sci. USA* 82, 1376-1380 (1985).
36. Coll, M., Frederick, Ch.A., Wang, A.H.-J. and Rich, A., *Proc. Nat. Acad. Sci. USA* 84, 8385-8389 (1987).
37. Zasedatelev, A.S., Zhuze, A.L., Zimmer, Ch., Grokhovsky, S.L., Tumanyan, V.G., Gursky, G.V. and Gottikh, B.P., *Dokl. Akad. Nauk USSR* 231, 1006-1009 (1976).
38. Gursky, G.V., Tumanyan, V.G., Zasedatelev, A.S., Zhuze, A.L., Grokhovsky, S.L. and Gottikh, B.P., *Nucleic Acid-Protein Recognition*, (Vogel, H.J., Ed.) Academic Press, New York, pp. 189-217 (1977).
39. Livshitz, M.A., Gursky, G.V., Zasedatelev, A.S. and Volkenstein, M.V., *Nucleic Acids Res.* 6, 2217-2237 (1979).
40. Gursky, G.V. and Zasedatelev, A.S., *Sov. Sci. Rev. D. Physicochem. Biol.* 5, 53-139 (1984).
41. Gursky, G.V., Surovaya, A.N., Kurochkin, A.S., Volkov, S.N., Chernov, B.K. and Kirpichnikov, M.P., submitted.
42. Ward, B., Rehffuss, R., Goodisman, J. and Dabrowiak, J.C., *Nucleic Acids Res.* 16, 1359-1369 (1988).
43. Vengerov, Yu.Yu., Semenov, T.E., Streltsov, S.A., Makarov, V.L., Khorlin, A.A. and Gursky, G.V., *J. Mol. Biol.* 184, 251-255 (1985).
44. Makarov, V.L., Streltsov, S.A., Vengerov, Y. Y., Khorlin, A.A. and Gursky, G.V., *Molekul. Biol. (USSR)* 17, 1089-2002 (1983).
45. Kramer, H., Niemoller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B. and Muller-Hill, B., *EMBO J.* 6, 1481-1491 (1987).
46. Brennan, R.G. and Matthews, B.W., *J. Biol. Chem.* 264, 1903-1906 (1989).

