Molecular Biology, Vol. 28, No. 5, Part 2, pp. 724-729, 1994

Interaction with DNA of a Synthetic Peptide Containing a Part of the DNA-Binding Domain of Transcription Activator v-Jun

S. L. Grokhovskii,[1] A. N. Surovaya,[1] A. L. Zhuze,[1] and G. V. Gurskii[1]

UDC 577.126

Translated from *Molekulyarnaya Biologiya,* Vol. 28, No. 5, pp. 1128-1136, September-October, 1994.

Original article submitted May 31, 1994.

Synthesis and interaction with DNA are described for a 26-residue peptide containing two copies of a fragment of the DNA-binding domain of the transcription activator v-Jun. The CD spectroscopy data showed the synthetic peptide to be in a random conformation in an aqueous solution, and partially in *a*-helical conformation in the presence of 20% trifluoroethanol. In 40% trifluoroethanol the relative content of *a*-helix increases to about 80%. The peptide was shown to form two types of complex with DNA. The first is saturated at one peptide molecule per six base pairs. With a further increase of the peptide/DNA molar ratio the binding becomes a cooperative process. The second complex is saturated at one peptide molecule per four DNA base pains. Jhe association constant for the first type of complex in the presence of 0.2 M NaCl was estimated at $1*10^{-5}$ M^{-1}. The peptide binds more strongly to poly[d(GC)]*poly(dC) and poly(dA)poly(dT) than to poly[d(GC)]*poly[d(GC)]*. We found that the DNA minor groove-binding antibiotic distamycin A competes effectively with the peptide for the binding to poly(dA)*poly(dT).

Key words: transcription activator; oncogene; v-Jun; DNA-binding domain; peptide

During the last years, eukaryotic transcription activators containing the "leucine zipper" are being investigated intensely. The examples of such proteins are the yeast transcription activator GCN4, proteins v-Jun, c-EBR, c-Fos, CREB, etc. They play a significant role in various biological processes, including oncogenesis.

Because of homology in amino acid sequences, all these proteins are similar in the three dimensional structure of their DNA-binding domains [2—8]. GCN 4 and v-Jun bind to DNA in the form of dimer [4, 5, 9, 10], whereas Fos does not form dimers in solution and binds to DNA as heterodimer with Jun [4, 6—8]. The structure of the dimer is stabilized by "adhesion" of *a*-helices and formation of a "leucine zipper" [3—9]. The 60-residue C-terminal fragment of GCN4 is known to bind with DNA as a dimer and recognize the same nucleotide sequence as the full-sized protein [10, 11]. The DNA-binding domain is located in the dimer N-termini, close to the leucine zipper, comprises approximately 25 residues, and is rich in lysine and aiginine [10, 14]. The DNA-binding domains of GCN4 and v-Jun are highly homologous. Both proteins in a dimer form recognize pseudosymmetrical and symmetrical nucleotide sequences in DNA: 5'-TGACTCA-3' (API site) and 5'-TGACGTCA-3' (CRE site) [4, 6—12]. It has been shown recently that a 34-residue peptide with two DNA-binding domains GCN4 connected with an S-S bond is able to recognize* the API site despite the absence of "leucine zipper" [14].

The aim of the present work was to synthesize a peptide which would contain two identical peptide fragments of the DNA-binding domain of the protein v-Jun. These two fragments were to be connected with a linker allowing both fragments to bind with two identical DNA sites simultaneously. Such a peptide in monomer form might have significant binding specificity. The main problem in constructing the peptide was that at the beginning of this work the structure of the peptide—DNA complex as well as the number of the residues in the v-June DNA-binding domain were not known precisely. We and other authors [14-18] had to search for the optimal peptide structure by gradual approximations. As the first step of the work, we synthesized a 26-residue peptide whose amino acid

[1]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 117984.

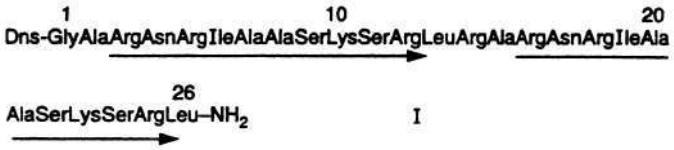

Fig. 1. Synthetic 26-residue peptide. Arrows mark the amino acid residues from the DNA-binding domain of the v-Jun protein (3-12 and 16-25). The direction of the arrow indicates the sequence N—C^{α}-C' in the peptide.

sequence is shown in Fig. 1. In this peptide, residues 3-12 and 16-25 found in the v-Jun DNA-binding domain are connected by a linker consisting of three amino acid residues.

MATERIALS AND METHODS

The peptide was synthesized manually by the solid-phase method on a phenylacetamidomethyl resin (PAM). *Tret*-butyloxycarbonyl (Boc) was used for a-amino group protection. For side chain protection the following groups were used: Tos (Arg), Bzl (Ser), 2-C1-Z (Lys). Boc-amino acids were used either as activated pentafluorophenyl or 1-hydroxybenzotriazolyl esters or as symmetrical anhydrides. Deprotection and simultaneous detachment of the peptide from the resin were performed with liquid HF in the presence of thioanisole and dimethyl sulfide. After gel filtration on Toyopearl HW-40 and preparative reversed-phase HPLC (column with C-18) we obtained about 300 mg of the peptide.

The structure of the 26-residue peptide was confirmed by mass spectrometry and amino acid analysis. The maximal peak in the mass spectrum, measured with an Electron (Sumy, the Ukraine) mass spectrometer *(m/z* = 3109.7) corresponded to a molecular ion of molecular mass 3112.68 Da. Amino acid analysis of the 26-residue peptide gave the following mole fractions:

Ala 5.8 (6); Arg 6.6 (7); Asp (Asn) 2.0 (2); Ile 1.8 (2); Leu 1.8 (2); Lys 2.2 (2); Ser 3.3 (4). Given in parentheses are the theoretical values corresponding to the chemical formula of the peptide.

The synthetic peptide was kept lyophilized and dissolved in 1 mM Na-cacodylate buffer (pH 7.0) immediately before the experiment. The peptide concentration was determined spectrophotometrically using the dansyl molar extinction coefficient e_{330} = 4300.

. The following preparations were used: bovine thymus DNA (e_{259} = 13,300, Sigma); synthetic polynucleotides poly(dA)*poly(dT) (e_{259} = 12,000), poly[d(GC)]*poly(dC) (e_{253} = 14,800), and poly[d(GC]* poly[d(GC)] (e_{260} = 16,800) from PL Biochemicals. The extinction values are calculated per 1 mole of DNA base pairs. Poly(dG)*poly(dC) was dissolved in 0.1 M NaOH and dialyzed for 24 h against the buffer in the presence of 1.0 mM EDTA. Prior to the experiments, all polynucleotide solutions were dialyzed twice against 1.0 mM Na-cacodylate buffer in the absence of EDTA. The CD spectra of the polynucleotides were compared with those from the literature data. The shape of the CD spectra of poly[d(G)]*poly(dC) changed gradually when the solutions were kept at 4°C. If spectra were distorted, the solution pH was changed (pH 11) and the dialysis procedure repeated.

Absorption spectra were measured using a Specord M40 spectrophotometer (Germany), CD spectra were measured with a Jasco-7200 dichrograph (Japan). Fluorimetric measurements were performed in the Aminco SPF-1000 Sc unit (USA).

RESULTS

CD spectra of peptide I in solution. Figure 2 demonstrates the CD spectra of the synthetic peptide in the absence (curve 1) and in the presence of trifluoroethanol (TFE) (curves 2 and 3). Typical of peptides in a random conformation, the CD spectrum of this peptide in the absence of trifluoroethanol displays a negative CD band at 197 nm. The negative band at 210-230 nm testifies to a portion of the peptide molecules being in an ordered conformation (*a*-helix or */3*-structure).

In the presence of 20% (vol.) TFE the amplitude of the negative CD band at 197 nm decreases; a characteristic two-dip negative band at 210 and 230 nm as well as a positive band at 190-195 nm appear. This means that peptide molecules in the presence of 20% TFE transit from a random conformation to the *a*-helical one. Both positive and negative CD bands increase significantly in the presence of 40% TFE, which testifies to the increase in the relative content of cc-helix. Using the procedure for estimating the helical content [19],

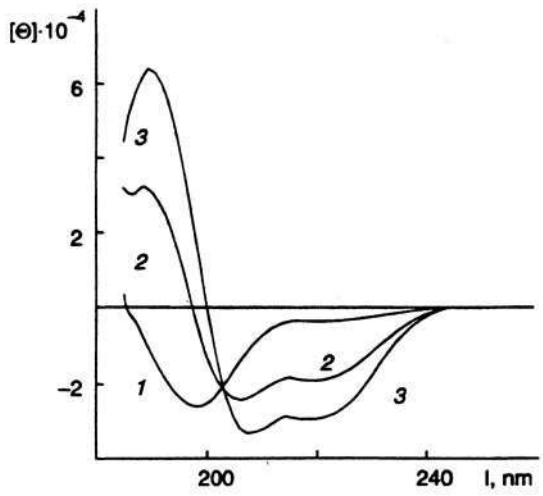

Fig. 2. CD spectra of the peptide in solution; [O] is the ellipticity expressed in grad·cm^2·$dmole^{-1}$ of amino acid residues. The content of TFE(%) was 0 (7), 20 *(2),* or 40 (3). The peptide concentration was 2 10^{-5} M in 1.0 mM cacodylate buffer, 0.1 M NH_4F, pH 7, 20°C.

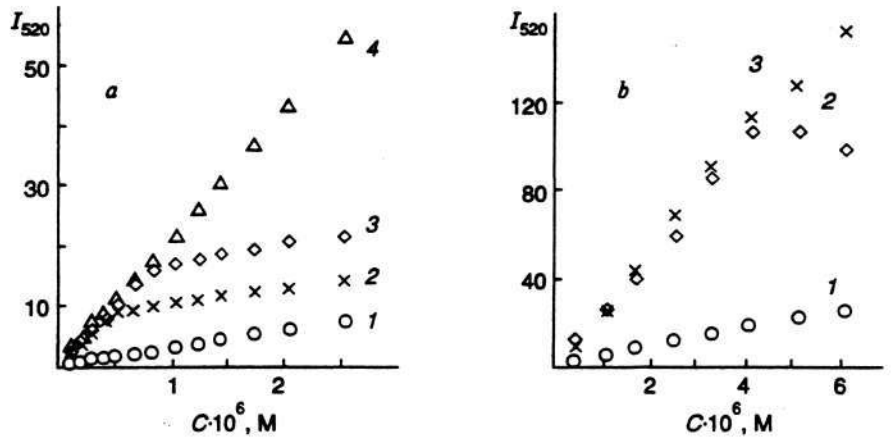

Fig. 3. Fluorimetric titration of DNA with the peptide. *C* is peptide concentration; I_{520} is the fluorescence intensity at 520 nm; the excitation wavelength was 380 nm; the slits for the excitation and the fluorescence were 10 nm. Panel *a*: *1*) free peptide; *2*–*4*) peptide in the presence of $1.7 \cdot 10^{-6}$, $3.4 \cdot 10^{-6}$, and $1.7 \cdot 10$ M^{-5} (bp), respectively; 0.001 M Na-cacodylate, 0.1 M NaCl, pH 7, 20°C. Panel *b*: *1*) free peptide; *2, 3*) peptide in the presence of $1.72 \cdot 10^{-5}$ and $8.6 \cdot 10^{-5}$ M (bp), respectively; 20% TFE, other conditions as in Fig. 3*a*.

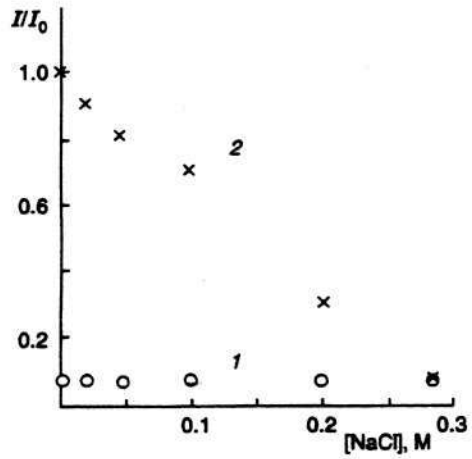

Fig. 4. Stability of the peptide–DNA complex versus NaCl concentration. DNA concentration is $2.7 \cdot 10^{-5}$ M, the peptide concentration is $1 \cdot 10^{-5}$ M. I_o is fluorescence intensity of the mixture of DNA with the peptide in a buffer without NaCl, *I* is that in the presence of NaCl.

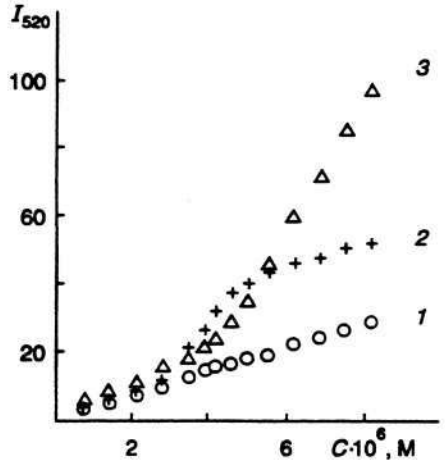

Fig. 5. Fluorimetric titration of DNA with the peptide at 0.2 M NaCl. *C* is the concentration of the peptide. I_{520} is the fluorescence intensity at 520 nm; the excitation wavelength is 360 nm. Slits for excitation and emission were 10 nm. *1*) free peptide; *2* and *3*) peptide in the presence of DNA, $8.86 \cdot 10^{-6}$ and $2.95 \cdot 10^{-5}$ M (bp), respectively. 0.001 M Na-cacodylate, 0.2 M NaCl, pH 7, 20°C.

we concluded that in the presence of 40% TFE the content of *a*-helix is 77% whereas in the presence of 20% TFE it is about 50%.

Interaction of synthetic peptide with DNA and synthetic polynucleotides. Because the peptide contained a fluorescent label (dansyl), we used fluorescence methods for registering its binding with DNA. Figure 3 shows the curves for peptide titration of solutions with different DNA concentrations in the absence of TFE (Fig. 3a) and in the presence of 20% TFE (Fig. 3*b*). One can see that the complex is saturated when one peptide molecule occupies four DNA base pairs both in the absence of TFE and in the presence of 20% TFE. Notably, the conformation of the peptide in solution depends strongly on the TFE content. The CD data prove that in the absence of TFE the peptide is preferentially in a random conformation. In the presence of 20% TFE the content of *a*-helical conformation is ~50%.

Figure *6a* shows the dependence of fluorescence polarization for the DNA—peptide complex versus the molar ratio of the peptide to the DNA base pairs. For comparison, Fig. 6*b* represents the curve of fluorimetric titration of DNA with peptide under the same conditions. Comparing these curves, one can see that the formation of the DNA—peptide complex of the second type is accompanied by a significant rise in the fluorescence polarization nearly to the maximum (41—42%). At occupancies higher than one peptide molecule per three base pairs, the solutions opalesce, and precipitation of DNA with the peptides is observed.

The increase in fluorescence polarization can be due to several causes, such as an increase in the complex molecular mass because of DNA aggregation; or tighter

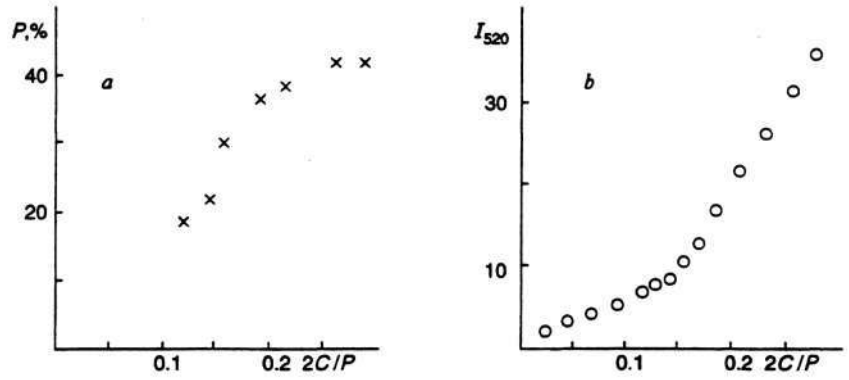

Fig. 6. *a:* Dependence of fluorescence polarization (P%) on *2C/P* (peptide per base pair), *b:* Dependence of fluorescence intensity at 520 nm (I_{520}) on *2C/P*. The excitation wavelength was 360 nm. Slits for excitation and emission were 10 nm for fluorescence intensity and 20 nm for fluorescence polarization assays. DNA concentration was $2.95*10^{-5}$ M (bp). 0.001 M Na-cacodylate, 0.2 M NaCl, pH 7, 20°C.

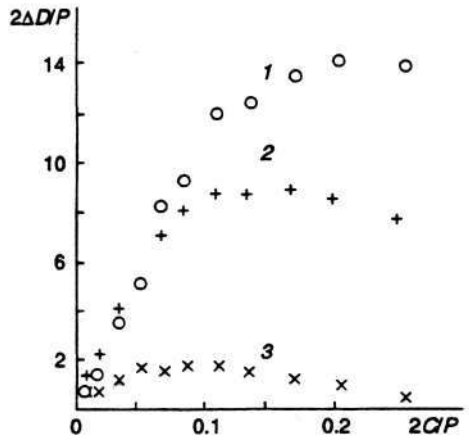

Fig. 7. Competition between the peptide and distamycin A for binding to poly(dA)*poly(dT); *1)* distamycin A titration of free poly(dA)*poly(dT); *2, 3)* distamycin A titration of poly(dA)*poly(dT) in the presence of $5.7\text{-}10^{-6}$ M and $1\ 10^{-5}$ M peptide, respectively. The poly(dA)*poly(dT) concentration was $2.6\ 10^{-5}$ M (bp). 2^*D/P* is dichroism at 320 nm counted per 1 cm of optical path and 1 mole base pairs, *2C/P* is the distamycin A quantity per base pair. 0.06 M NaCl, 0.001 M Na-cacodylate, pH 7, 20°C.

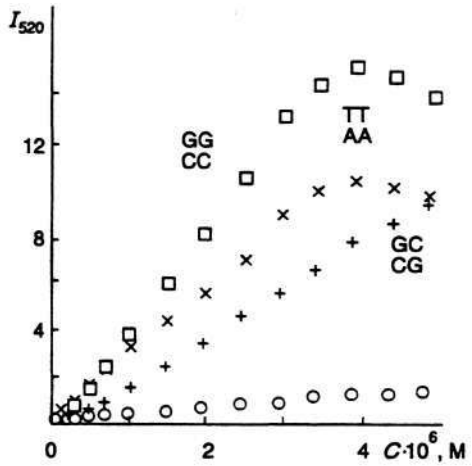

Fig. 8. Fluorimetric titration with the peptide of various polynucleotides: poly[d(GC)]*poly(dC), poly(dA)*poly(dT), and poly[d(GC)]*poly[d(GC)]. The polynucleotide concentration in all cases was 1.4 10 M^{-5} in base pairs. The circles in the figure indicate the dependence of fluorescence intensity on the peptide concentration (C) in the absence of the polynucleotide. 0.1 M NaCl, 0.001 M Na-cacodylate, pH 7, 20°C

fixation of the fluorescence label upon formation of the oligomer complex, which could be accompanied by a change of the DNA conformation.

Binding of the peptide depends strongly on the solution ionic strength, because of a large number of lysine and arginine residues within the peptide. With the increase in NaCl concentration, the amount of the peptide bound with DNA decreases (Fig. 4). At 0.15 NaCl, approximately half of the bound peptide dissociates. At 0.2 M NaCl, not only the binding constant but also the shape of the titration curve change. Figures 3 and 5 show the curves for the fluorimetric titration of DNA with the peptide at 0.1 M and 0.2 M NaCl, respectively. One can see a marked difference in the shapes of the titration curves obtained at 0.2 M and 0.1 M NaCl. In the presence of 0.2 M NaCl the peptide forms two types of complexes with DNA. The first type of complex, probably monomer, has low fluorescence intensity per mole of the peptide. The fluorescence polarization changes from 9% for the free peptide to 19% for the peptide bound to DNA as a monomer. This binding is a noncooperative process. The association constant for this type of complex in the presence of 0.2 M NaCl is approximately 10 M. The monomer type of complex is saturated when one peptide molecule occupies six base pairs (Fig. 5).

With a further increase of the peptide to DNA ratio the binding becomes a cooperative process and is attended by a rise in fluorescence intensity. This type of complex we called oligomer, as the cooperativity is probably due to the interaction between the peptide molecules bound with the neighboring sites on DNA.

Binding of the peptide with DNA is testified by the fact that in the presence of the peptide the binding of other ligands, e.g., DNA minor-groove-binding anti-

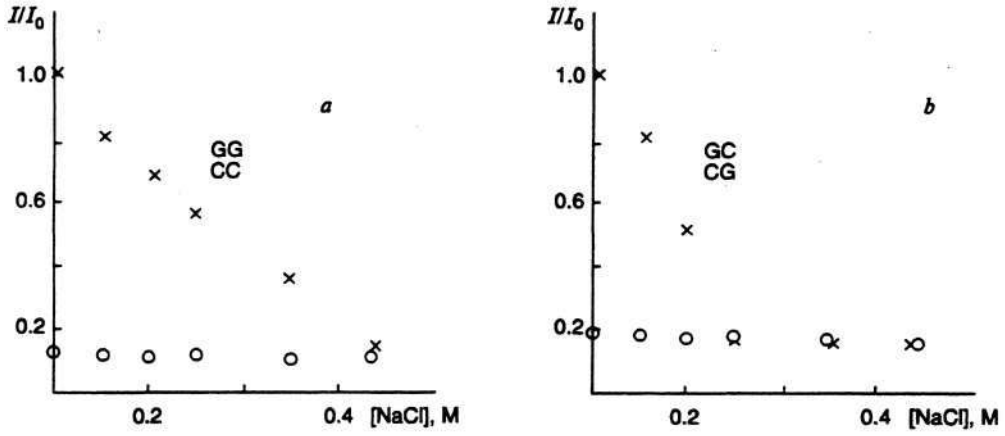

Fig.9. Stability of complex of the peptide with poly[d(G)]*poly(dC) *(a)* and poly[d(GC)]*poly[d(GC] *(b)* versus NaCl concentration. The polymer concentration is 1.4 10^{-5} bp. Peptide concentration is 5.0 10^{-6} M. I_0 and I are fluorescence intensities of the peptide-polynucleotide mixture in the absence and in the presence of NaCl, respectively. Circles represent the dependence of fluorescence intensity on NaCl concentration for the free peptide. Conditions are given in Fig. 8.

biotic distamycin A, decreases. Figure 7 presents the curves of distamycin A titration of the free poly(dA)* poly(dT) (curve 1) and that in the presence of the peptide (curves 2 and 3). One can see that in the presence of the peptide, the amount of the antibiotic bound is significantly lower than that in the case of the free poly(dA)*poly(dT). These data testify to competitive binding with poly(dA)*poly(dT) between the peptide and distamycin A.

To elucidate the nucleotide sequences in DNA to which the peptide binds selectively, we obtained the DNase I footprinting diagrams. These results appear in the next paper [15]. Another way to study the specificity of the peptide binding is to compare the curves for peptide titration of various synthetic polynucleotides. Figure 8 represents the curves for titrations of poly[d(G)]* poly(dC), poly(dA)*poly(dT), and poly[d(GC)]* poly[d(GC)] with the peptide. In these experiments, the same polynucleotide concentration was used, 1.4 10 M (bp). One can see that the peptide binding depends strongly on the nucleotide sequence. The peptide binds tighter to poly(dG)*poly(dC) and poly(dA)*poly(dT) than to poly[d(GC)]* poly[d(GC)]. This is also supported by the fact that the complexes of the peptide with poly[d(G)]*poly(dC) appeared to be more stable against NaCl than those with poly[d(GC)]*poly[d(GC] (Fig. 9). Thus, the peptide can discriminate homo- and hetero(GC) sequences.

DISCUSSION

Our experiments showed that the synthetic 26-residue peptide containing two identical fragments from the DNA-binding domain of the v-Jun protein can bind to DNA and form two types of complex: a monomer type saturated at the peptide/DNA ratio equal to one peptide molecule per six base pairs, and an oligomer one with a stoichiometry of one peptide molecule per four base pairs.

According to the CD spectroscopy data, the peptide in an aqueous solution is in a random conformation in the absence of trifluoroethanol (TFE), and assumes a partly *a*-helical conformation in the presence of 20% TFE (~50% of *a*-helix). Despite the difference in the peptide conformation in solution in the absence and in the presence of 20% TFE, the shape of the titration curves as well as the complex stoichiomethry (1 peptide molecule per 4 base pairs) are the same. Within the free GCN4 and v-Jun, the DNA-binding domains are but slightly ordered [27—30]. At normal intracellular concentrations, GCN4 dimers are practically absent. However, both dimerization and stabilization of the leucine zipper and DNA-binding domain are induced upon binding to DNA [27-30], According to the literature data, a similar transition from a random to ordered conformation takes place at binding to DNA in the peptides from the DNA-binding domain of GCN4 [14, 17, 20]. We obtained similar results for the peptide I analogs as well as for other specific DNA-binding peptides [15, 20].

The complex stability and the shape of the titration curves depend on the solution ionic strength. In the presence of 0.2 M NaCl the binding is a cooperative process.

It was shown that the peptide binding to DNA hinders the binding of distamycin A. At high peptide concentrations, distamycin A practically does not bind to DNA. The competition between the peptide and distamycin A for the DNA-binding sites may testify to the occupation of the minor groove by both the peptide

and the antibiotic. It may also be that, while binding to the major groove, the peptide changes the DNA conformation so as to hinder distamycin A binding in the minor groove. It should be noted that an analog of distamycin-related antibiotic netropsin with altered binding specificity as well as the C-terminal fragment of the GCN4 transcription activator can bind simultaneously to the same DNA site in the minor and major grooves, respectively [26].

The binding constant of the peptide markedly depends on the nucleotide sequence in the binding site. The peptide binds more strongly to poly[d(G)]*poly(dC) and poly(dA)*poly(dT) than to poly[d(GC)]* poly[d(GC)j. The footprinting diagrams of the peptide analog, having the residues Ala and Leu at positions 2 and 13 replaced with cysteine residues with a blocking Acm group on the side chain, show that nucleotide sequences S'-TGY$(R)_n$TGY-S' (where Y is A or C, R is any nucleotide, n = 0, 1) are found in the preferential binding sites for the peptide. The nucleotide sequence 5'-TGA-3' is present in the specific binding site API for the transcription activator v-Jun. Thus, the recognition of the sequence 5'-TGY$(R)_n$TGY by the peptide is probably due to the presence in the synthetic peptide of two copies of a fragment from a DNA-binding domain of the protein v-Jun [15]. These results are consistent with the data concerning the specific interactions between peptide residues 3—12 (or 16—25) and DNA.

According to the X-ray analysis for the CGN4—DNA complex, the following residues participate in specific contacts with DNA bases within the major groove: Asn-235, Arg-243 and Ala-238 and 239 [24]. Within the synthetic peptide, these residues are located at positions 4, 12, 7, and 8, as well as 17, 25, 20, and 21, respectively. Recently the photocrosslinking method was used to demonstrate that the $C_b$atom of Ala 238 is close to the thymine methyl group in the complex of GCN4 with DNA [25].

Thus peptide I conserves all the residues which in the complex of the full-sized protein with DNA contribute to the specificity of the interaction with DNA bases. However, it lacks some residues interacting with the DNA sugar-phosphate backbone within the crystalline GCN4-DNA complex. The residues which constitute the leucine zipper are absent as well. For the peptides from the DNA-binding domains GCN4 and v-Jun, of particular interest is their ability to recognize the specific nucleotide sequences not only when the peptide is part of the protein globule but also when the peptide is isolated and is in a random conformation in solution. The search of the minimal-length peptides which are able to recognize the specific nucleotide sequences on DNA is of great interest for constructing new specific DNA-binding ligands.

The work was supported by the Russian Fund of Basic Research and the program Frontiers in Genetics.

REFERENCES

1. P. F. Johnson and S. L. McKnight, *Annu. Rev. Biochem.,* **58**, 799-839 (1989).
2. T. K. Kerppola and T. Cunan, *Curr. Opin. Struct. Biol, 2,* 109-115(1991).
3. W. H. Landchulz, P. F. Johnson, and S. L. McKnight, *Science,* **240,** 1759-1764 (1988).
4. T. Kouzarides and E. Ziff, *Nature,* **336,** 646-651 (1988).
5. W. H. Landchulz, P. F. Johnson, and S. L. McKnight, *Science,* **243,** 1681-1688 (1989).
6. C. R. Vinson, P. B. Sigler, and S. L. McKnight, *Science,* **246,** 911-916(1989).
7. R. Gents, F. J. Rauscher, C. Abate, and T. Curran, *Science,* **243.** 1695-1699 (1989).
8. Yu. Nakabeppu and D. Nathans, *EMBO J.,* 8, 3833-3841 (1989).
9. E. K. O'Shea, J. D. Klemm, P. S. Kim, and T. Alber, *Science,* **254,** 539-544 (1991).
10. I. A. Hope and K. Stnihl, *Cell,* 46, 885-894 (1986).
11. K. Stnihl, *Cell,* **50,** 841-846 (1987).
12. A. Oliphant, Ch. J. Brandl, and K. Stnihl, *Mol. Cell. Biol,* **9**, 2944-2949 (1989).
13. T. J. Boss, F. J. Ramscher, T. Curran, and P. K. Vogt, *Oncogene,* 4, 123-126 (1989).
14. R. V. Talanian, S. L. McKnight, and P. S. Kim, *Science,* **249,** 769-773 (1990).
15. S.L. Grokhovskii, A.N. Surovaya, A.L. Zhuze, and G.V. Gurskii, *Mol Biol,* **28,** 1137-1148 (1994).
16. M. G. Oakley and P. B. Dervan, *Science,* **248,** 847-850 (1990).
17. K. T. O'Neil, R. H. Hoess, and W. F. DeGrado, *Science,* **249,** 774-778 (1990).
18. Ch. Park, J. L. Campbell, and W. A. Goddard, *Proc. Natl Acad. Sci. USA,* **89,** 9094-9096 (1992).
19. N. Sreerama and R. W. Woody, *Anal. Biochem.,* **191,** 110-118 (1993).
20. S.L. Grokhovsky, A.N. Surovaya, R.V. Brussov, B.K. Chernov, N.Yu. Sidorova, and G.V. Gursky, *J. Biomol. Struct. Dyn.,* **8**, 989-1025 (1991).
21. M. L. Kopka, C. Joon, D. Goodsell, P. Pjura, and R. E. Dickerson, *Proc. Natl. Acad. Sci. USA,* **82,** 1376-1380 (1985).
22. M. Coll, C. A. Frederick, A. H-J. Wang, and A. Rich, *Proc. Natl Acad. Sci. USA,* **84,** 8385-8389 (1987).
23. R. E. Klevit, D. E. Wemmer, and B. R. Reid, *Biochemistry,* **25**, 3296-3303 (1986).
24. T. E. EUenberger, Ch. J. Brandy, K. Stnihl, and S. C. Harrison, *Cell,* **71,** 1223-1237 (1992).
25. E. E. Blatter, Y. W. Ebright, and R. H. Ebright, *Nature,* 359, 650-652 (1992).
26. M. G. Oakley M., Mrksich, and P. B. Dervan, *Biochemistry,* **31,** 10969-10975 (1992).
27. K. T. O'Neil, J. D. Shuman, C. Ampe, and W. F. DeGrado, *Biochemistry,* **30,** 9030-9034 (1991).
28. M. A. Weiss, T. E. EUenberger, C. R. Wobbe, J. P. Lee, S. C. Harrison, and K. Stnihl, *Nature,* **347,** 575-578 (1990).
29. V. Saudek, H. S. Pasley, T. Gibson, H. Gausepohl, R. Frank, and A Pastore, *Biochemistry,* 30, 1310-1317 (1991).
30. L. Patel, C Abate, and T. Curran, *Nature,341,* 572-574 (1990).