Molecular Biology, Vol. 31, No. 4, 1997, pp. 623–630. Translated from Molekulyarnaya Biologiya, Vol. 31, No. 4, 1997, pp. 733–740.

UDC 577.113.3; 57.087.1

Computer Conformation Analysis of TATA-Box Sequences in Eukaryotic Promoters

M. P. Ponomarenko[1], Yu. V. Ponomarenko[1], A. E. Kel'[1], N. A. Kolchanov[1], H. Karas[2], E. Wingender[2], and H. Sklenar[3]

[1] *Institute of Cytology and Genetics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
E-mail: pon@bionet.nsc.ru
[2] *Gesellschaft für Biotechnologische Forschung mbH (GBF), Mascheroder Weg 1, D-38124 Braunschweig, Germany*
[3] *Max Delbrück for Molecular Medicine, Robert-Rossle-Str. 10, D-13122 Berlin, Germany*

Received December 16, 1996

Abstract—We have developed a method for revealing local conformation properties of binding sites for transcription factors, based on conformation indices of B-sheet DNA hexanucleotide duplexes and computer system SITEVIDEO. Analysis of sequences with TATA box has revealed the promoter regions with certain conformation indices of B-sheet DNA significantly changed as compared with arbitrary sequences: TATA-box DNA has a lower angle of helical twist, smaller distance between neighbor base pairs along the DNA helix axis, wider minor and narrower major DNA grooves.

Key words: TATA-box, computer analysis, B-DNA, transcription

INTRODUCTION

The genes transcribed by RNA polymerase II are the most common in eukaryotic genome [1]. Approximately 80% promoters of these genes carry TATA boxes more or less corresponding to the consensus sequence TATAAAAA [2]. Transcription is initiated by assembly of the general transcription complex at the promoter region, which is tens of base pairs long and is located upstream to the transcription start [1]. In its turn, assembly of the general transcription complex on the TATA-box promoters starts with recognition of the TATA box by the TATA-binding protein (TBP) [1].

The important role of TATA box in formation of the transcription complex has focused considerable attention on it. X-ray analysis of the DNA–protein complexes indicates that yeast and plant TBP binds the TATA box at the minor groove bending the DNA by 80° towards the major groove [3, 4]. The DNA bend is provided by four phenylalanine residues intercalated between the neighbor DNA bases at the TATA-box termini.

The process of TBP–TATA binding has the following stages [5, 6]: nucleosome displacement, site-nonspecific TBP binding to the DNA, TBP diffusion along the DNA, site-specific TBP binding to the TATA box, isomerization of the TBP–TATA-box complex, and formation of stable DNA–protein complex.

DNA bending within the complex with TBP is sequence-dependent and correlates to stability of the complex and TBP capacity for transcription [7]. Hence, the presented data suggest that DNA conformation within the complex with TBP is important for transcription initiation. Nonetheless, the role of DNA conformation in TBP–TATA-box interaction remains unclear.

One approach to solving this problem is based on studying the local DNA conformation of TATA-containing sequences. Studies of the effect of the nucleotide sequence on B-DNA conformation followed the pioneer X-ray analysis of dodecamers [8, 9] resulting a set of rules describing certain local conformation properties of B-DNA as a function of nucleotide sequence [10–14]. One of approaches to study the relationship between local DNA conformation and nucleotide context is based on the methods of conformation analysis, molecular mechanics and dynamics [15]. These methods allow calculation of low-energy conformation of DNA duplexes for various nucleotide sequences [16, 17]. Using this approach, Sklenar [18]

0026-8933/97/3104-0623$18.00 © 1997 МАИК Наука /Interperiodica Publishing

Table 1. Conformation indices for certain hexanucleotide B-DNA duplexes [18]

k	Index	AGGGAT*	TTTTTT*	AAGCCT*
1	Twist, deg	34.14	38.90	38.81
2	Rise, Å	3.97	3.19	3.73
3	Tip, deg	–2.44	0	0.59
4	Inclination, deg	0.03	0	1.90
5	Bending angle, deg	2.44	0	1.99
6	Width of the minor groove (w)**, Å	4.22	5.16	4.89
7	Depth of the minor groove (d)**, Å	9.46	9.09	9.40
8	Width of the major groove (W)**, Å	17.36	9.70	14.44
9	Depth of the major groove (D)**, Å	8.11	8.91	9.12

* Nucleotide sequences of the main strand of the duplex (5'–3').
** Values for the 3rd base pair in the hexanucleotide.

has built a library of B-DNA conformation indices for all 4096 possible hexanucleotide duplexes. Structural study of yeast TATA boxes implementing all these data [19] has demonstrated profile similarity of certain DNA conformation indices in the TATA box region.

Here we present a method for revealing local conformation properties of functional DNA sites. It is based on computer system SITEVIDEO [20, 21] dealing with the site nucleotide sequences and conformation indices of hexanucleotide duplexes [18].

We analyzed TATA-containing nucleotide sequences of vertebrate, invertebrate, and yeast promoters. B-DNA in the TATA box region features smaller angle of helical twist and smaller distance between neighbor base pairs along the axis of the DNA helix as compared with arbitrary sequences. In addition the TATA boxes featured widening of the minor groove and narrowing of the major groove as compared with arbitrary sequences. Investigation of the *E. coli* TATA-containing promoter regions (Pribnow box) has also demonstrated decreased distance between the neighbor base pairs and width of the major groove and increased width of the minor groove. Still, by contrast to eukaryotic TATA boxes, the Pribnow box DNA has a greater angle of helical twist as compared with arbitrary sequences.

MATERIALS AND METHODS

Conformation indices of B-DNA. We used the library of B-DNA conformation indices for all possible hexanucleotide duplexes [18]. There are the following indices defined according to the standard nomenclature [22, 23]: twist, the twist angle between neighbor base pairs relative to the axis of DNA helix; rise, the distance between neighbor base pairs along the axis of DNA helix; twist angles of a base pair plane relative to the long (tip) and short (inclination) axes of the base pair (the long axis goes through atoms C6 of pyrimidine and C8 of purine, the short axis is perpendicular to the long one); bending angle of the DNA axis; and width and depth of the major and minor grooves.

The library values for twist, rise, tip, inclination, and bending angle of the DNA double helix are given for the central base pair in the duplexes. Width and depth of the minor and major grooves are given for the 3rd and 4th duplex base pairs in the library. We averaged these values to produce the indices describing the duplex center. Table 1 illustrates conformation indices for a number of duplexes according to the library [18]. Table 2 presents the minimum, maximum, and mean values as well as the dispersion for each index calculated for all 4096 hexanucleotide duplexes in the library. One can see significant variability of the conformation indices relative to their mean values. This indicates the influence of the nucleotide sequences on local conformation of B-DNA.

Nucleotide sequences of TATA-containing promoters. We studied the nucleotide sequences presented in Table 3. We considered four samples of the TATA-containing promoter sequences in vertebrates, invertebrates, yeast, and *E. coli* and a sample of arbitrary sequences with even nucleotide frequencies. Yeast and *E. coli* promoters were extracted from EMBL Data Library (release 42) by keywords "promoter" and "primary transcript." Promoters of vertebrates and invertebrates were taken from EPD (release 45). Accounting rather variable position of TATA box in eukaryotic promoter [2], we searched them using either consensus sequence (in case of yeast and *E. coli*) or weight matrix (in case of vertebrates and invertebrates). All TATA-containing sequences were

Table 2. Mean values and ranges of conformation indices for certain hexanucleotide B-DNA duplexes calculated on the basis of the library [18]

k	Index	min	max	Mean
1	Twist, deg	30.48	42.54	36.14 ± 3.91
2	Rise, Å	2.99	4.83	3.53 ± 0.33
3	Tip, deg	–8.12	12.15	1.34 ± 2.96
4	Inclination, deg	–4.31	4.31	0.0 ± 1.44
5	Bending angle, deg	0.0	12.15	3.03 ± 1.86
6	Width of the minor groove (w)**, Å	3.81	7.50	5.18 ± 0.76
7	Depth of the minor groove (d)**, Å	7.75	10.37	9.00 ± 0.44
8	Width of the major groove (W)**, Å	8.44	18.52	13.51 ± 1.63
9	Depth of the major groove (D)**, Å	7.28	10.53	8.92 ± 0.56

Note: See Table 1.

Table 3. Sequences carrying TATA box

Taxon	Sequence property		TATA box	
	number	length, bp	nucleotide	search method
Vertebrates	486	70	1	weight matrix
Invertebrates	158	70	1	″
Yeast	75	70	1	consensus
E. coli	135	70	1	″
Arbitrary sequences	500	70	none	none

70 bp long. First position of the TATA box in each sequence was designated as number one. A sample of TATA-containing vertebrate promoters is given below:

```
EMBL index      -31                                1                                39
PRI : HSTUBB2   TTCCGAGACTAGCGGAGGCGGGCAGGGAGGGTATAATAAGCGTTGGCGGACGGTCGGTTGTAGCACTCTGC
VRT : GGAC01    TTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCG
VRT : GGMYHE    GAGGTGGCTGCTACGTATGCAAATCAGAGCCTATAAAGGACCTTAGGGTCAGTGTGTCTTGTCCTTCTT
.....
.....
.....
VRT : GGOV01    GCTAACAGGGGGAGTGTCCGAAAGGGTACTGTATAATATCACCAAGGACTCAGAGAATCTGTTCAGGTTCA
PRI : HSAPOA2   TCACCTGACAGGGGGTGGGTAAACAGACAGGTATAATAGCCCCTTCCTCTCCAGCCAGGGCAGGCACAGAC
ROD : RNALBPR   GTTAATGATCTACAGTTATTGGTTAGAGAAGTATAATTAGAGCGAGTTTCTCTGCACACAGACCACCTTTC
```

The sequences in each sample were aligned relative to the first position of the TATA box. Thus arranged promoter sample was designated as $\{S^+\}$, i.e., containing TATA box, while a sample of arbitrary sequences was designated as $\{S^-\}$, i.e., lacking a TATA box.

Building and analysis of a B-DNA conformation index profile. Let us consider nucleotide sequence $S_n = \{ s_1^n \ldots s_i^n \ldots s_{70}^n \}$ (n = 1, …, N) of a sample containing N sequences. Let us focus on kth index describing DNA conformation (Table 1). Let us consider ith position of a hexanucleotide window sliding within the range of a given sequence (i = 1, …, 65) and assign it a value of the conformation index $X_{ki}(S_n)$ corresponding to the hexanucleotide in the current window. For instance, let us consider rise (k = 2) and a hexanucleotide AAGCCT at position i, then this position is assigned $X_{2i}(S_n)$ = 3.73. In case of hexanucleotide TTTTTT $X_{2i}(S_n)$ = 3.19 (Table 1). One can calcu-

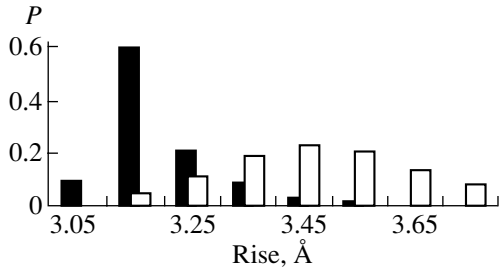

Fig. 1. Distribution of distance between neighbor base pairs along the DNA axis averaged for the [–2; 8] region for vertebrate promoter and arbitrary sequences. Here and below: filled and empty bars represent promoters and arbitrary sequences, respectively.

late the mean value for thus built profile for *k*th conformation index $X_{ki}(S_n)$ within a given region [*a*, *b*] (*a*, *b* = 1, …, 65) in a sequence S_n:

$$X_{k,a,b}(S_n) = \sum_{i=a}^{b} X_{ki}(S_n)/(b-a+1). \qquad (1)$$

In this case we can evaluate the distribution of $X_{k,a,b}(S_n)$ for all sequences in the considered sample. Such distribution $X_{2,-2,8}(S_n)$ for rise (k = 2) within the [–2; 8] region in the samples of TATA-containing vertebrate promoters and arbitrary sequences is exemplified in Fig. 1. One can see that the distribution $X_{2,-2,8}(S_n^+)$ for the promoters is shifted leftward relative to that for arbitrary sequences. Comparison of these distributions leads to the conclusion that in general TATA-containing promoters have decreased rise at the region [–2; 8].

In order to assess significance of the difference between the two $X_{k,a,b}$ distributions and its discriminating capacity for separating the $\{S^+\}$ and $\{S^-\}$ samples we used the approach previously implemented in the SITEVIDEO computer system [20, 21]. It relies on processing a lot statistical indices of the compared distributions: extent of their difference from the normal distribution, degree of their overlapping, difference between the minimum, maximum, and mean values and dispersions, stability of the distribution indices to formation of arbitrary subsamples of $\{S^+\}$ and $\{S^-\}$ (a total of 14 statistical indices were considered). Based on these indices we calculated the difference in $X_{k,a,b}$ for the samples $\{S^+\}$ and $\{S^-\}$ described by $U(X_{k,a,b})$. This value is called utility in the decision theory [24] and conforms to:

(1) $-1 < U(X) < 1$;

(2) if $U(X) > 0$ then the X index is significant for discrimination between $\{S^+\}$ and $\{S^-\}$;

(3) if $U(X) > U(Y)$ than X index is more significant for discrimination between $\{S^+\}$ and $\{S^-\}$ than Y index.

For the distributions presented in Fig. 1, $U(X)$ = 0.887, which is close to the top utility (unity). This indicates a significant decrease ($\alpha < 10^{-40}$) in rise relative to arbitrary sequences. When studying any sample of TATA-containing sequences, we tested all possible conformation indices k = 1, …, 9. We considered all possible regions [*a*, *b*] from 6 to 70 bp within the sequence for a given k. The number of such regions was (70 – 6 + 2)(70 – 6 + 1)/2 = 2145. For a given region [a, b] we generated $X_{k,a,b}(S_n)$ distribution for the samples $\{S^+\}$ and $\{S^-\}$. Condition (2) was used to select the regions [*a*, *b*] with $X_{k,a,b}$ significant for discriminating samples $\{S^+\}$ and $\{S^-\}$. At the next stage condition (3) was used to select the region [*a*, *b*] with the highest difference by the *k*th index between TATA-containing promoters and arbitrary sequences.

RESULTS AND DISCUSSION

The results of studying the nucleotide sequences of TATA-containing promoters in vertebrates, invertebrates, yeast, and *E. coli* are summarized in Table 4.

Four conformation indices significantly different from arbitrary sequences were revealed in the vertebrate TATA-containing sequences:

(1) Rise averaged within the region [–2; 8] is lower for the promoters (3.18 ± 0.09 Å) than for arbitrary sequences (3.48 ± 0.17 Å). Utility of this index for discriminating samples $\{S^+\}$ and $\{S^-\}$ U = 0.887. Reliability of the difference between distribution of this index in the promoter and arbitrary sequences samples $\alpha < 10^{-40}$ (Fig. 1).

(2) Twist averaged within the region [–4; 5] is lower for the promoters (34.48° ± 1.23°) than for arbitrary sequences (36.02° ± 1.91°). Utility of this index U = 0.632. Figure 2a visualizes the significant leftward shift of this index distribution for the promoter sample as compared with arbitrary sequences ($\alpha < 10^{-12}$).

(3) Width of the minor groove averaged within the region [–2; 6] is higher for the promoters (5.80 ± 0.29 Å) than for arbitrary sequences (5.15 ± 0.58 Å). Figure 3a presents the distribution of this index for the samples $\{S^+\}$ and $\{S^-\}$. Reliability of the difference between these two distributions $\alpha < 10^{-13}$. Utility of this index for discriminating samples $\{S^+\}$ and $\{S^-\}$ U = 0.692.

(4) Width of the major groove averaged within the region [–2; 6] is lower for the promoters (11.98 ± 0.4 Å) than for arbitrary sequences (13.37 ± 1.3 Å).

The data presented in Table 4 and Figs. 2 and 3 demonstrate similar conformation properties of B-DNA in TATA box regions of invertebrate and yeast promoters: decreased rise and twist, increased width of the minor groove, and decreased width of the major groove. As concerns TATA-containing sequences in *E. coli*, they also feature decreased rise and twist,

Table 4. Significant conformation indices of TATA-containing promoter B-DNA

Taxon	Index	[*a*; *b*]	$U(X_{k,a,b})$	$X_{k,a,b}$	Mean value		*P*	α
					promoters	arbitrary sequences		
Vertebrates	Rise	[–2; 8]	0.887	$X_{2,-2,8}$	3.18 ± 0.09	3.48 ± 0.17	<	$<10^{-40}$
	Twist	[–4; 5]	0.632	$X_{1,-4,5}$	34.48 ± 1.23	36.02 ± 1.91	<	$<10^{-12}$
	Width of the minor groove	[–2; 6]	0.692	$X_{6,-2,6}$	5.80 ± 0.29	5.15 ± 0.58	>	$<10^{-13}$
	Width of the major groove	[–2; 6]	0.671	$X_{8,-2,6}$	11.98 ± 0.4	13.37 ± 1.3	<	$<10^{-13}$
Invertebrates	Rise	[–5; 5]	0.837	$X_{2,-5,5}$	3.18 ± 0.07	3.49 ± 0.14	<	$<10^{-40}$
	Twist	[–4; 5]	0.664	$X_{1,-4,5}$	33.75 ± 0.67	36.02 ± 1.38	<	$<10^{-14}$
	Width of the minor groove	[–4; 8]	0.834	$X_{6,-4,8}$	6.06 ± 0.26	5.16 ± 0.47	>	$<10^{-14}$
	Width of the major groove	[–2; 10]	0.976	$X_{8,-2,10}$	11.79 ± 0.5	13.33 ± 1.0	<	$<10^{-13}$
Yeast	Rise	[–21; 14]	0.979	$X_{2,-21,14}$	3.39 ± 0.05	3.50 ± 0.06	<	$<10^{-7}$
	Twist	[–4; 5]	0.940	$X_{1,-4,5}$	34.01 ± 0.65	36.02 ± 1.50	<	$<10^{-7}$
	Width of the minor groove	[–4; 7]	0.857	$X_{6,-4,7}$	5.98 ± 0.25	5.15 ± 0.43	>	$<10^{-6}$
	Width of the major groove	[–10; 9]	0.958	$X_{8,-10,9}$	12.38 ± 0.4	13.41 ± 0.7	<	$<10^{-7}$
E. coli	Rise	[–29; 32]	0.571	$X_{2,-29,32}$	3.47 ± 0.05	3.50 ± 0.04	<	$<10^{-8}$
	Twist	[–36; 17]	0.637	$X_{1,-36,17}$	36.36 ± 0.42	36.05 ± 0.35	>	$<10^{-8}$
	Width of the minor groove	[–1; 12]	0.562	$X_{6,-1,12}$	5.36 ± 0.4	5.14 ± 0.4	>	$<10^{-7}$
	Width of the major groove	[–23; 11]	0.644	$X_{8,-23,11}$	12.98 ± 0.5	13.41 ± 0.5	<	$<10^{-8}$

Note: [*a*; *b*] significant region; $U(X_{k,a,b})$, utility of U; $X_{k,a,b}$, property; *P*, reference to arbitrary sequences: the mean index for promoters is smaller (<) or greater (>) as compared with arbitrary sequences. For units of measure see Table 1.

* The mean value ± standard deviation for promoter and arbitrary sequences.

increased width of the minor groove, and decreased width of the major groove.

Approximation of B-DNA conformation from its nucleotide sequence was based on the library of hexanucleotide duplex conformation indices [18] as substantiated by correspondence between the conformation indices presented in it and X-ray data (Fig. 4).

For instance, the calculated and experimental twist for dodecamers CGCGAATTCGCG [9] and CGATCGATCG [25] reliably corresponds (coefficient of linear correlation $r = 0.906$ and 0.946, respectively, $\alpha < 0.01$) (Fig. 4a). For 60 hexanucleotides with twist determined by X-ray analysis [9, 25–32] the coefficient of linear correlation to the indices calculated on basis of the library [18] $r = 0.361$ ($\alpha < 0.01$) (Fig. 4b).

Above we have shown that the eukaryotic TATA boxes feature decreased twist within the region [–4; 5]. This agrees with the X-ray analysis data. The mean twist of DNA (TATAAAAG) within the plant TBP–TATA-box complex was 31.7° [4]. The corresponding values for yeast TBP–TATA-box (TATATAAA) complex were even lower and ranged from 3.05° to 22.85° [3], which is significantly below the standard B-DNA twist, 36.0° [15].

Interestingly, the putative DNA sites of binding to the nucleosome have greater twist as compared with arbitrary sequences (Fig. 5). Analysis of 39 nucleosome binding sites (141 bp [33]) have demonstrated that the mean B-DNA twist of these sites (36.50° ± 0.28°) is reliably greater than that of arbitrary sequences (36.05° ± 0.22°). Hence, TATA boxes with small angle of helical twist disfavor nucleosome binding and favor release of the nucleosome in the course of TBP–TATA binding, which agrees with the experimental data [34].

The obtained data suggest that a wider minor groove within the TATA box as compared with arbitrary sequences is a significant conformation property of promoter DNA in vertebrates, invertebrates, yeast, and *E. coli*. Apparently, this property is related to TBP interaction with the minor groove of the TATA box DNA [3, 4]. This observation also agrees with the X-ray data on TBP complexes with TATA-containing DNA duplexes [3, 4] demonstrating that width of the minor DNA groove in the region of TATA box ranges from 7.4 to 9.9 Å. At the same time, width of the minor groove in the regular B-DNA helix and Dickerson–Drew dodecamer is 4.2–6.7 and 3.8 Å, respectively [3, 9]. Comparison of these values demonstrates that the DNA in the region of TBP–TATA-box complex has wider minor groove as compared to the standard B-DNA double helix. Apparently, wide minor groove is important for specific TBP–DNA binding.

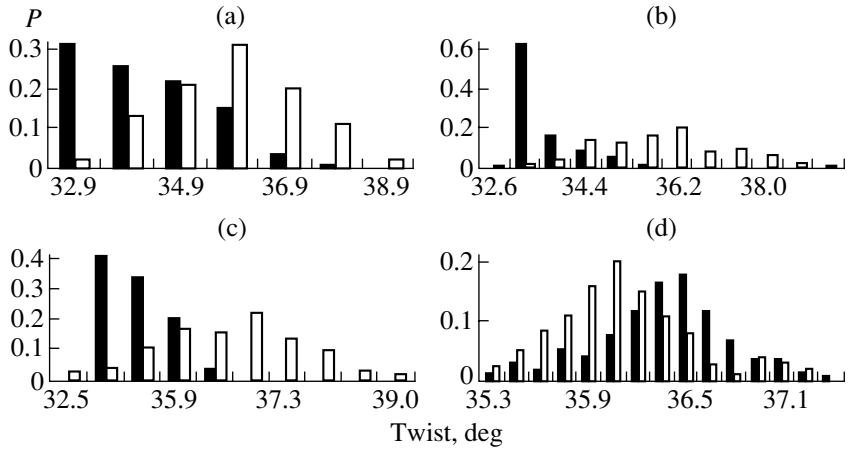

Fig. 2. Distribution of B-DNA twist for promoter and arbitrary sequences: (a) vertebrates (averaged for the [–4; 5] region); (b) invertebrates (averaged for the [–4; 5] region); (c) yeast (averaged for the [–4; 5] region); (d) *E. coli* (averaged for the [–36; 17] region).

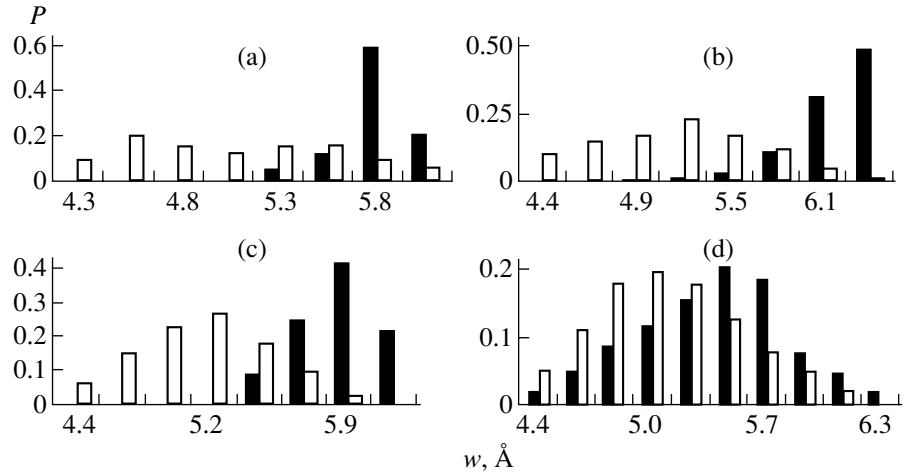

Fig. 3. Distribution of width of B-DNA minor groove for promoter and arbitrary sequences: (a), vertebrates (averaged for the [–2; 6] region); (b), invertebrates (averaged for the [–4; 8] region); (c) yeast (averaged for the [–4; 7] region); (d) *E. coli* (averaged for the [–29; 32] region).

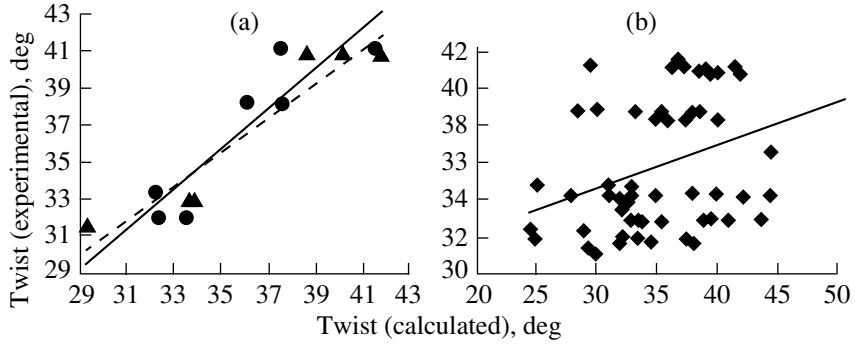

Fig. 4. Correlation between the calculated [18] and experimental indices of B-DNA: (a) twist for CGCGAATTCGCG (circles, continuous line) [9] ($r = 0.906$, $\alpha < 0.01$) and CGATCGATCG (triangles, dotted line) [25] ($r = 0.949$, $\alpha < 0.01$); (b) twist for dodecamer B-DNA [12, 25–32] ($r = 0.361$, $\alpha < 0.01$).

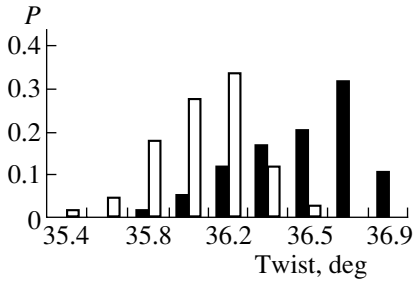

Fig. 5. Distribution of twist for nucleosome-binding sites [33] and arbitrary sequences.

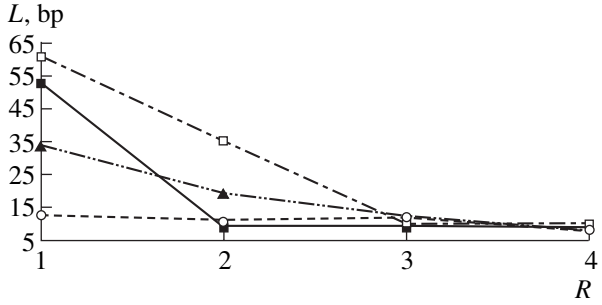

Fig. 6. Relationship between size of significant region (*L*) and taxonomic range of the organism (*R*) for significant conformation indices of TATA-box B-DNA. (*1*) *E. coli*; (*2*) yeast; (*3*) invertebrates; (*4*) vertebrates; ○-----○, width of the minor groove; ▲– – –▲, width of the major groove; □–––––□, rise; ■———■, twist.

It was interesting to consider the size of the promoter regions where significant conformation indices were revealed. In case of width of the major groove, the significant region was 8, 12, 19, and 34 bp for vertebrates, invertebrates, yeast, and *E. coli*, respectively (Fig. 6). Hence, there is a clear trend for decreasing the significant region in the taxonomic series *E. coli* ⟶ yeast ⟶ invertebrates ⟶ vertebrates. A similar relationship between size of the significant region and taxonomic range is true for width of the minor groove and rise (Fig. 6). As concerns twist, both localization of the significant region relative to the TATA box ([–4; 5]) and its size (9 bp) are uniform in eukaryotes but it is significantly smaller as compared to *E. coli* (53 bp). It is possible that these relationships reflect increasing complexity of the transcription complex in the above series.

In this work we propose a method for revealing conformation properties of DNA sites for binding transcription factors based on the conformation index library for B-DNA hexanucleotide duplexes [18] and computer system SITEVIDEO [20, 21]. Analysis of TATA-containing sequences has demonstrated smaller twist and rise, wider minor and narrower major grooves in the TATA-box DNA as compared with arbitrary sequences. These results agree with the experimental data on the structural and functional organization of TATA boxes and point to specific local conformation of TATA-containing promoter sequences.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 95-04-12469 and 97-04-4970), Russian State program Human Genome, and Siberian Branch of Russian Academy of Sciences.

REFERENCES

1. Latchman, D.S., *Eukaryotic Transcription Factors*, London: Academic, 1995, pp. 47–59.
2. Bucher, P., *J. Mol. Biol.*, 1990, vol. 212, pp. 563–578.
3. Kim, Y., Geiger, J.H., Hahn, S., and Sigler, P.B., *Nature*, 1993, vol. 365, pp. 512–520.
4. Kim, J.L., Nikolov, D.B., and Burley, S.K., *Nature*, 1993, vol. 365, pp. 520–527.
5. Coleman, R.A. and Pugh, B.F., *J. Biol. Chem.*, 1995, vol. 270, pp. 13850–13859.
6. Coleman R.A., Taggart K.P., Benjamin L.R., and Pugh, B.F., *J. Biol. Chem.*, 1995, vol. 270, pp. 13842–13849.
7. Starr, D.B., Hoopes, B.C., and Hawley, D.K., *J. Mol. Biol.*, 1995, vol. 250, pp. 434–446.
8. Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., and Dickerson, R.E., *Nature*, 1980, vol. 287, pp. 755–758.
9. Dickerson, T.D. and Drew, H.R., *J. Mol. Biol.*, 1981, vol. 149, pp. 761–769.
10. Calladine, C.R., *J. Mol. Biol.*, 1982, vol. 161, pp. 343–357.
11. Dickerson, R.E., *J. Mol. Biol.*, 1983, vol. 166, pp. 419–425.
12. Calladine, C.R. and Drew, H.R., *J. Mol. Biol.*, 1984, vol. 178, pp. 773–785.
13. Zhurkin, V.B., *Mol. Biol.*, 1983, vol. 17, pp. 622–638.
14. Ul'yanov, R.B. and Zhurkin, V.B., *Mol. Biol.*, 1984, vol. 18, pp. 1664–1685.
15. Neidle, S., *DNA Structure and Recognition*, New York: IRL, 1994.
16. Lavery, R., Sklenar, H., Zakrzewska, K., and Pullman, B., *J. Biomol. Struct. Dyn.*, 1986, vol. 3, pp. 989–1014.
17. Lavery, R., Zakrzewska, K., and Sklenar, H., *Comp. Phys. Commun.*, 1995, vol. 91, pp. 135–158.
18. Sklenar, H., *Proc. Int. Workshop on Computational Analysis of Eukaryotic Transcriptional Regulatory Elements*, GBF, 1996, pp. 44–47.
19. Karas, H., Knuppel, R., Schulz, W., Sklenar, H., and Wingender, E., *Comput. Applic. Biosci.*, 1996, vol. 12, pp. 441–446.
20. Kel, A.E., Ponomarenko, M.P., Likhachev, E., Orlov, Yu.L., Ischenko, I.V., Milanesi, L., and Kolchanov, N.A., *Comput. Applic. Biosci.*, 1993, vol. 9, pp. 617–627.
21. Ponomarenko, M.P., Kel, A.E., Orlov, Yu.L., Benjukh, D.N., Ischenko, I.V., Bockhonov, V.B., Likhachev, E.A., and Kolchanov, N.A., *Computer Analysis of Genetic Macromolecules: Structure, Function and Evolution*, Kolcha-

nov, N.A. and Lim, H., Eds., Singapore: World Sci., 1994, pp. 35–65.

22. Lavery, R. and Sklenar, H., *J. Biomol. Struct. Dynam.*, 1988, vol. 6, pp. 63–91.
23. EMBO Workshop, *EMBO J.*, 1989, vol. 8, pp. 1–4.
24. Fishburn, P.C., *Utility Theory for Decision Making*, New York: Wiley, 1970.
25. Grzeskowiak, K., Yaanagi, K., Prive, G.G., and Dickerson, R.E., *J. Biol. Chem.*, 1991, vol. 266, pp. 8861–8883.
26. Yoon, C., Prive, G.G., Goodsell, D.S., and Dickerson, R.E., *Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, pp. 6332–6336.
27. Nelson, H.C.M., Finch, J.T., Luisi, B.F., Klug, A., *Nature*, 1987, vol. 330, pp. 221–226.
28. Webster, G.D., Sanderson, M.R., Skelly, J.V., Neidle, S., Swann, P.F., Li, B.F., and Tickle, I.J., *Proc. Natl. Acad. Sci. USA*, 1990, vol. 87, pp. 6693–6697.
29. Yuan, H., Quintana, J., and Dickerson, R.E., *Biochemistry*, 1992, vol. 31, pp. 8009–8021.
30. Leonard, G.A. and Hunter, W.N., *J. Mol. Biol.*, 1993, vol. 234, pp. 198–208.
31. Heinemann, U., Alings, C., and Bansal, M., *EMBO J.*, 1992, vol. 11, pp. 1931–1939.
32. Prive, G.G., Yanagi, K., and Dickerson, R.E., *J. Mol. Biol.*, 1991, vol. 217, pp. 177–199.
33. Ioshikhes, I. and Trifonov, E.N., *Nucleic Acids Res.*, 1993, vol. 21, pp. 4857–4859.
34. Godde, J.S., Nakatani, Y., and Wolffe, A.P., *Nucleic Acids Res.*, 1995, vol. 23, pp. 4557–4564.