Acta Biochimica Polonica

Vol. 45 No. 1/1998

221-231

QUARTERLY

Binding of SPXK- and APXK-peptide motifs to AT-rich DNA. Experimental and theoretical studies[o*]

Jan Brzeski[1], Tomasz Grycuk[2], Andrzej W. Lipkowski[3], Witold Rudnicki[2], Bogdan Lesyng[2,4] and Andrzej Jerzmanowski [1,5✉]

[1]*Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5A, 02-106 Warsaw, Poland,* [2]*Department of Biophysics, Warsaw University, F. Żwirki i S. Wigury 93, 02-089 Warsaw, Poland,* [3]*Medical Research Center, Polish Academy of Sciences, A. Pawińskiego 5, 02-106 Warsaw, Industrial Chemistry Research Institute, J. Rydygiera 8, 01-793 Warsaw, Poland,* [4]*Interdisciplinary Center for Mathematical and Computational Modeling, Warsaw University, A. Pawińskiego 5A, 02-106 Warsaw, Poland, and* [5]*Laboratory of Plant Molecular Biology, Institute of Experimental Plant Biology, Warsaw University, A. Pawińskiego 5A, 02-106 Warsaw, Poland*

Received: 30 October, 1997

Key words: SPXK-peptides, DNA binding

The binding properties of the SPXK- and APXK-type peptides to the AT-rich DNA fragments of different length were studied by measuring the competition of peptides with Hoechst 33258 dye for DNA binding and by the gel shift assay analysis. In parallel to the experimental studies, molecular modeling techniques were used to analyze possible binding modes of the SPXZ and APXK motifs to the AT-rich DNA. The results of the competition measurements and gel shift assays suggest that serine at the *i–1* position (*i* is proline) can be replaced by alanine without affecting the binding properties of the motif. Thus, the presence of the conserved serine in this motif in many DNA-binding proteins is probably not dictated by structural requirements. Based on the results of molecular modeling studies we propose that the binding mode of the SPXK-type motifs to the AT-rich DNA resembles closely that between the N-terminal arm of the homeodomain and DNA. This model confirms that serine in the SPXK motifs is not essential for the DNA binding. The model also indicates that if X in the motif is glutamic acid, this residue is probably protonated in the complex with DNA.

[o]This work was supported by the State Committee for Scientific Research grants: 6PO4A 053 08 for T. Grycuk and 6PO4A 029 13 for A. Jerzmanowski and J. Brzeski.

[*]Fee for the MSI software was covered from the State Committee for Scientific Research, a grant 2 P3O2 029 07.

[✉]Corresponding author: Andrzej Jerzmanowski, phone: (48-22) 659 6072 ext. 3234; fax: (48-22) 658 4636; e-mail: andyj@ibbrain.ibb.waw.pl.

Abbreviations: BOP, benzotriazol-1-yloxy-tris(dimethylamino)phosphonium hexafluorophosphate; HOBt, 1-hydroxybenzotriazole; MS-FAB, mass spectrometry/fast bombardment; PB, Poisson-Boltzman (model).

Many sequence-specific DNA binding proteins, such as transcription factors and restriction enzymes recognize and bind their cognate sequences using domains that interact with the major groove of DNA (for review see [1, 2]). In contrast, the minor groove binders are less specific. Proteins that utilize the minor groove binding sites of the AT-rich DNA are histones, high mobility group proteins and other structural components of chromatin [3, 4]. Several drugs, such as Hoechst 33258, netropsin and distamycin also bind into the minor groove [5–7]. Other proteins, like the GATA factor [8] or those containing the homeodomain are known to contact DNA *via* both the major and the minor groove [9]. For many DNA-binding proteins neither the structure nor the binding mode of the minor groove binding motifs are well known. This also concerns the so called 'SPXK' motif found originally in histones and subsequently in many other DNA-binding proteins [3, 10]. The consensus sequence of this motif is S/TPXZ, where X is any, and Z a basic residue. It has been proposed that the 'SPXK' fragment binds the minor groove of the AT-rich DNA and recognizes the same structural features as are recognized by distamycin [3]. In addition to binding the DNA, the 'SPXZ' motif is a specific target for the mitotic cyclin–kinase complex responsible for a reversible phosphorylation of histone H1 that occurs universally in proliferating eukaryotic cells [11]. The biological significance of this phosphorylation is not well understood. It is possible that the phosphorylable serine in the minor groove-binding motif may serve as a "switch" regulating the binding properties of the whole protein.

There are numerous variants of the consensus 'S/TPXZ' motif in the N- and C-terminal domains of different H1 histones. The histones differ also in the number and localization of these motifs with respect to their characteristic central globular domain. Histone H1 has been shown to regulate the transcriptional activity of certain classes of genes [12–14]. In order to study the importance of the specific variants of H1 in regulating the cellular processes dependent on chromatin structure we have been using the system of transgenic plants enabling partial or complete replacement of histone H1 variants *in vivo* [15]. Thus, it was important to learn the role of the 'SPXZ' motifs in some of these variants in their interaction with DNA. Because of the putative regulatory role of serine phosphorylation we were also interested in defining the contribution of this residue to the overall binding properties of this motif.

In the present work we studied the *in vitro* binding of the synthetic 'SPXK'-type peptides to the AT-rich DNA fragments of different length. We showed that serine at the *i–1* (*i* is proline) position, being a target for the mitotic phosphorylation, can be replaced by alanine without any substantial effect on the binding properties.

In parallel to experimental studies, the molecular modeling techniques were used to analyze possible binding modes of the "SPXK' motifs to the AT-rich regions of B-DNA. The interpretation of the experimentally observed specific interactions between biomolecules is usually based on mezoscopic models, see e.g. [16, 17]. One of them, the Poisson-Boltzmann (PB) model, allows to determine of the molecular electrostatic potential provided the discrete atomic charges are known [17, 18]. The model also takes into account a continuous mobile charge distribution representing a cloud of ions surrounding the molecular objects in solution at a given temperature. We used the determined electrostatic potential to compute the mean electrostatic interaction energy between the DNA molecule and the studied peptides.

Experimental and computational procedures

Synthesis of peptides and DNA. All peptides were synthesized by a Fmoc strategy on

solid state using BOP with HOBt for the coupling reaction [19]. Final deprotections were accomplished using 95% trifluoroacetic acid. The purifications were performed on Sephadex G-15 in 0.1 M acetic acid, followed by preparative HPLC. The resulting pure peptides gave correct amino-acid analyses and molecular ion analyses in MS-FAB.

DNA oligonucleotide was purchased from Center of Macromolecular Research Polish Academy of Sciences (Łódź, Poland). To obtain double-stranded DNA, the self-complementary oligonucleotide was heated to 95°C and slowly cooled to room temperature.

DNA manipulations. The pUC 19 plasmid with inserted *Xenopus laevis* oocyte-type 5S RNA gene with its native flanks was kindly provided by J. Gottesfeld (Scripps Institute, San Diego, CA, U.S.A.). To obtain an insert with the 5S RNA gene, the plasmid DNA was cut with the *Hin*dIII restriction enzyme (Gibco BRL) and the resulting fragments were isolated with Gene-Clean Kit (Bio101) after separation of DNA by agarose gel electrophoresis.

Determination of the binding constants. All measurements were done in 40 mM Hepes (Sigma) at pH 7.0 and 60 mM NaCl. To the solution of peptide (8×10^{-6} M), DNA oligonucleotide (7.5×10^{-7} M) and Hoechst 33258 dye (Sigma) were added, the latter to a final concentration ranging from 1×10^{-7} to 8×10^{-6} M. After 1 h of incubation at room temperature the fluorescence of the DNA-dye complex was measured at λ_{exc} of 354 nm and λ_{ems} of 450 nm, using a Shimadzu RF 5000 fluorimeter.

Gel shift assays. The DNA fragment was dephosphorylated with calf thymus phosphatase (Gibco BRL) and labelled with [^{32}P]ATP using T4 polynucleotide kinase (Gibco BRL) according to the manufacturer's protocol. DNA (5×10^{-5} M calculated as DNA phosphate) was mixed with peptides at concentrations varied from 5×10^{-4} M to 5×10^{-3} M in 40 mM Hepes, at pH 7.0 and 60 mM NaCl. Following 1 h incubation the samples were electrophoresed in 1% agarose in 0.045 M Tris/borate, pH 8.0 and 1 mM EDTA. After electrophoresis DNA was blotted on a nylon membrane (MagnaGraph, MSI) and exposed to Phosphore Storage Screen (Molecular Dynamics). Detection of radioactive DNA was done using a Phosphor Imager (Molecular Dynamics).

Computational procedures. The theoretical studies of the interactions between peptides and oligodeoxyribonucleotides consisted of four stages:

- 1) On the basis of the available crystallographic and/or NMR data and by applying the molecular graphics techniques, the selected binding modes (approximate complex structures) of the interacting peptides and the model DNA duplex in the B form were created.
- 2) The geometries of the complexes were refined using molecular mechanics (MM) and molecular dynamics (MD) techniques, in particular using energy minimization methods and short MD relaxation runs (for review of the molecular modeling methods, see e.g. [20, 21]).
- 3) Based on the PB model and using the optimized geometries, the calculations of the electrostatic contribution to the free energy of binding were carried out with a DelPhi, MSI molecular modeling package. A finite difference approximation method discretizing the dielectric constant and the charge distributions on a three-dimensional cubic grid was applied. The energy of binding was calculated as the difference between the total energy of the whole system and the energies of the non-interacting objects. A smoothing procedure which improves precision and stability of the numerical solution of the PB equation was applied [22].
- 4) The optimized structures and the PB free electrostatic energies were analyzed. Based on this analysis and results of former theoretical studies [3, 23] the most probable binding modes were proposed and visualized.

RESULTS

The primary structure of peptides, oligodeoxynucleotides and the natural DNA fragment used in the experiments and in the molecular modeling studies

A schematic structure of the peptides, oligodeoxynucleotides and DNA fragments used in the studies are shown in Table 1. The KSPEKK peptide, denoted hereafter as the SPEK peptide, is the same one as the six residue motif occurring in the N-terminal domain of histone H1-2 from the plant *Arabidopsis thaliana* (position 22–26). The KAPEKK peptide, denoted hereafter as the APEK peptide, differs from the SPEK one in position *i–1*, containing alanine in the location of serine. The KRPRTA peptide, denoted hereafter as the RPRT one, is identical with the five-member N-terminal homeodomain fragment of the *Drosophila melanogaster* engrailed protein [9]. The oligodeoxynucleotide (GCGCAATTGCGC) used in these studies was the same as the dodecamer co-crystallized with netropsin [6]. Table 2 contains the list of peptides used in our theoretical studies.

The natural DNA fragment used in the gel-shift experiments was the 669-bp *Xenopus laevis* DNA, consisting of the 120-bp oocyte-type 5S RNA gene, accompanied by the 360-bp 5′- and the 189-bp 3′ flanks. The flanks which are 76% A + T, comprise the internally repetitious sequence, most of which can be represented as:

```
      C          AA    C   G
CAA GTTTTCAA   TTTGAA TTTTCT
      A          GC    A   T
```

Competition of peptides with Hoechst 33258 for DNA binding

The binding of Hoechst 33258 to DNA can be measured by monitoring fluorescence of the dye–DNA complex. In earlier studies it has been demonstrated that the SPXK-type peptides compete with Hoechst 33258 for the same binding sites in the minor groove of DNA [3, 22]. In order to determine the strength of the peptide binding to DNA we titrated the dodecamer DNA containing a well defined binding site for the dye, in the absence or presence of the peptide, with increasing amounts of Hoechst 33258. Titration with

Table 1. Oligopeptides and nucleic acids used in the experimental systems

No.	Sequence	Abbreviation	Origin and comments
Oligopeptides			
1	KSPEKK	SPEK	*A. thaliana H1*
2	KAPEKK	APEK	*A. thaliana H1*
3	KRPRTA	RPRT	*engrailed* homeodomain
Nucleic acids			
1	CGCGAATTCGCG		single binding site oligonucleotide
2	Genbank No. M10850		*X. laevis* oocyte-type 5S rRNA gene unit

Table 2. List of the oligopeptides used in the theoretical studies

No.	Sequence	Abbreviation	Comments
1	KSPKKK	SPKK	consensus motif
2	KAPKKK	APKK	S to A mutation of the consensus motif
3	KRPRTA	RPRT	a reference homeodomain
4	KSPEKK	SPEK	peptide used in experiments
5	KAPEKK	APEK	peptide used in experiments

the dye was done at a fixed peptide concentration of 1.0×10^{-4} M. The concentration of oligodeoxynucleotide was 1.9×10^{-4} M (calculated per DNA phosphate). Figure 1 demonstrates that the specific binding of Hoechst 33258 to the dodecamer DNA, as measured by fluorescence of the DNA–dye complexes, is inhibited both by the presence of the SPEK and the APEK peptide.

Using the Wilkinson method [24] the K_d for the dye and the K_{app} (the apparent dissociation constant) for the dye in the presence of peptides were determined basing on the data shown in Fig. 1. K_d for the SPEK and APEK

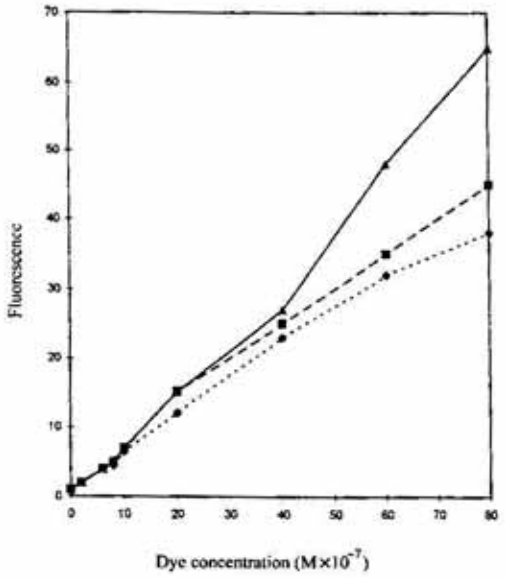

Figure 1. Binding of Hoechst 33258 dye (▲) to DNA.

Dependence of fluorescence intensity on the concentration of Hoechst 33258 in the presence of SPEK (●) and APEK (◆) peptides. Fluorescence from the DNA–Hoechst 33258 complex was measured at 450 nm, excited at 354 nm. The concentration of DNA was fixed at 7.5×10^{-7} M and the concentration of the dye was changed from 1×10^{-7} to 8×10^{-6} M.

peptides was calculated using the following equation 1:

$$K_d = K_{app}3D\ K_d\ (\text{dye})\ (1 + I/K_d\ (\text{peptide})),$$

where K_{app} is the apparent dissociation constant for dye binding in the presence of a competitive peptide. The free energy of binding (Table 3) was calculated as

$$\Delta G = -RT\ln K_d. \quad (2)$$

Table 3. Free binding energies of ligands to the DNA dodecamer (GCGCAATTGCGC)

Ligand	ΔG (kcal/mol)
Hoechst 33258	-9.16
SPEK	-9.16
APEK	-9.69

Binding of the peptides to natural DNA fragment rich in short oligo(dA)·oligo(dT) tracts

The results cited above show that the exchange of Ser for Ala does not diminish the competition of the SPXK-type peptide with Hoechst 33258 for the DNA binding (Fig.1). These results were obtained using a short DNA fragment containing a single binding site for the dye. It was interesting to check if the two peptides interact similarly with longer DNA that contains multiple potential bindings sites for the SPXK-type motifs. We used the natural DNA fragment of *X. laevis* comprising the repeating unit of the oocyte-type 5S RNA gene that consists of the 120-bp gene and the adjacent flanks. The fragment has multiple oligo(dA) · oligo(dT) tracts (see the sequence above) with a narrow minor groove which constitute an ideal spatial pockets for minor groove binding agents. Such tracts were shown to be preferred binding sites for both Hoechst dye and the SPXK motif [23]. In our experiment the end-labelled DNA was titrated with the peptide and the resulting complexes were subjected to electrophoresis in agarose gel (Fig. 2). Both the SPEK and APEK peptides bind efficiently to the tested DNA fragment as demonstrated by the upward shifts of DNA on the gels. The pattern of binding is similar for the APEK and SPEK peptides, and consistent with formation of multiple binding products and aggregates of different size. This is what one would expect for a

relatively long DNA fragment, virtually saturated with the binding sites. Also shown on Fig. 2 are the results of titration of the same AT-rich DNA with the RPRT peptide, representing the minor groove binding motif of the N-terminal arm of the homeodomain. The peptide does bind to DNA as indicated by the upward shifts of the DNA band at increased peptide concentration, although the complexes formed are less heterogeneous and less prone to aggregation than in the case of the APEK and SPEK peptides.

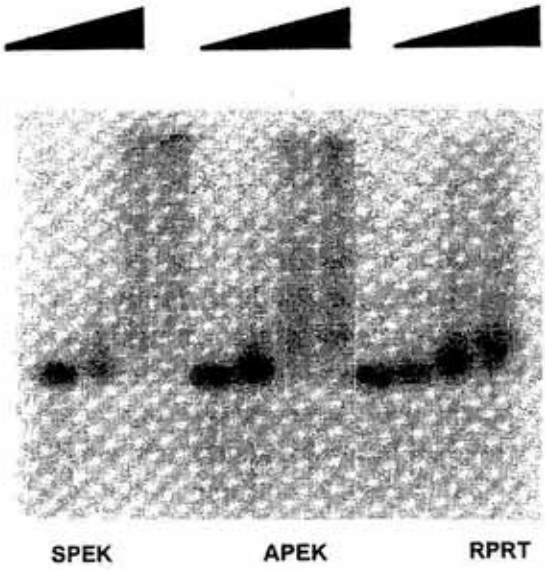

Figure 2. Gel shift assay of peptide–DNA complexes.

The labelled DNA fragment (5×10^{-5} M) comprising the repeating unit of oocyte-type 5S RNA gene of *X. laevis* was mixed with increasing amounts (from 5×10^{-4} M to 5×10^{-3} M) of indicated peptides and after 1 h incubation electrophoresed on 1% agarose gel. The position of radioactive DNA bands was visualized with Phosphor Imager. The first lane in each series represents DNA without the peptide. Lanes 2, 3 and 4 in each series correspond to 5×10^{-4}, 10^{-3} and 5×10^{-3} M of the peptide, respectively.

Theoretical studies of the interactions between the SPXK-type peptides and AT-rich DNA

In order to generate the approximate structures of the SPXK- and APXK-DNA complexes, we followed at first the procedure described by Suzuki *et al.* [3, 25] for their netropsin-type model. We encountered similar problems as the authors of this model. We observed steric clashes of the peptide side chains with the backbone phosphates upon docking the peptides into the minor groove of the d(AT) duplex. In the original study the peptides had to be moved outwards from the minor groove, and even then the tensions in the molecular structure of the complex remained. As the tensions in the molecular structures increase, their free energy increases too. Thus, the netropsin-type binding mode becomes doubtful. We therefore considered other binding modes.

To find an alternative way of building the model, a DNA binding unit containing the SPXK peptide pattern was searched for in a PIR database of the protein sequences. Such pattern was found in a homeodomain N-terminal arm, known to bind into the minor groove of the AT-rich DNA [9]. There is no crystallographic data for this particular molecular system. However, assuming the existence of a functional analogy between the homeodomains and the similarity of their three dimensional structures one can use a crystallographic structure determined for another homeodomain, as a reference. For this purpose we have chosen the system consisting of a *Drosophila melanogaster* engrailed homeodomain protein, complexed with DNA, for which the crystallographic data were available [9]. The protein part of this complex contained a suitable analogue for the SPXK fragment.

It was interesting to check how well the SPXK sequences, including the peptide sequences used in our experiments, fit into the sequences of the N-terminal arms of the homeodomains. After searching over two hundred available homeodomain sequences with the PATTERNS program from the GCG package we found that while the glutamic acid existing in the SPEK and APEK peptides does not occur (except in one case) in the homeodomains, the SPKK motifs fit reasonably well into the homeodomain sequences. This may suggest that glutamic acid changes properties

of the peptides of the SPXK family, reducing their affinity for binding into the minor groove of DNA. In fact, glutamic acid, being a typical hydrogen bond acceptor, is more likely to bind *via* the carbonyl group into the major groove of DNA. In the minor groove there are hydrogen bond acceptors. It is possible, however, that in a strong electric field generated by DNA, the pK of glutamic acid is shifted, so that its carboxyl group is protonated and becomes the hydrogen bond donor. In this form glutamic acid could bind into the minor groove.

The crystallographic structure of the engrailed homeodomain–DNA complex does not contain the typical SPXK motif. The homology analysis shows that in the SPKK position, with proline as a reference residue, there is a RPRT sequence. This sequence forms a β-turn and binds into the minor DNA groove. We used this structure to create models of the complexes. A suitable fragment of the homeodomain was exchanged for the peptide sequences, while the original backbone and the orientation of the side chains remained unchanged. The structures were refined using the energy minimization techniques.

The modeling scheme described above was applied to the peptide motifs listed in Table 2 and the crystallographic structure of the homeodomain fragment. Figure 3 presents the complex of the homeodomain SPKK peptide with a fragment of DNA directly engaged in the binding. The hydrogen bonds which stabilize the complex structure of homeodomain are conserved in this system (Table 4).

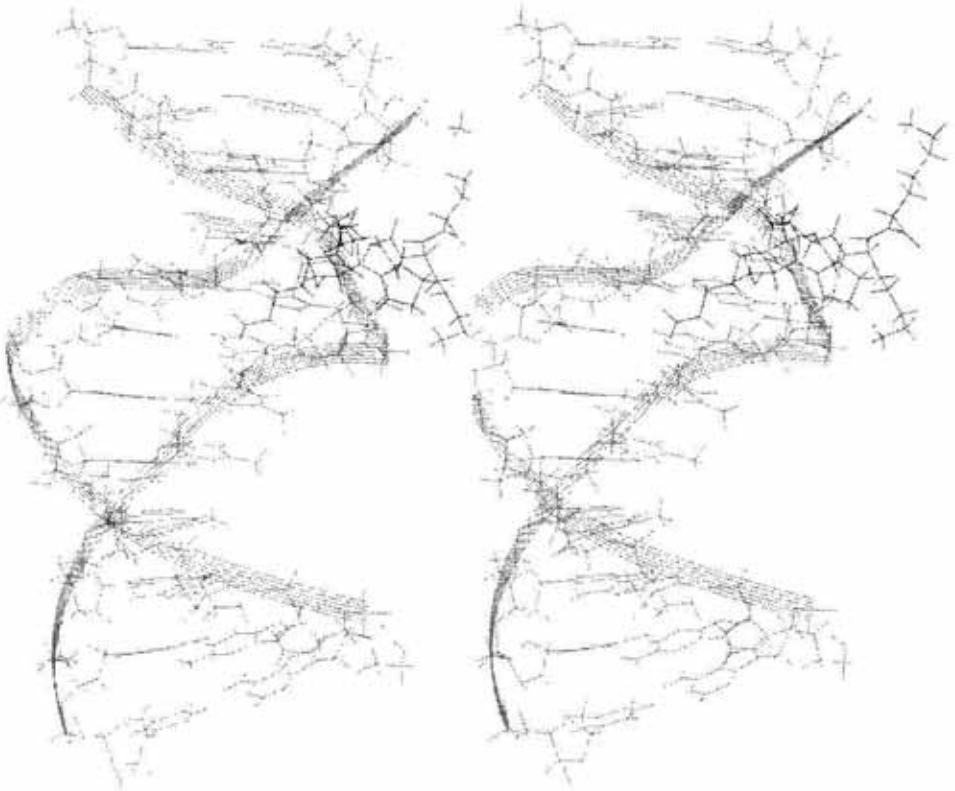

Figure 3. Stereo diagram of a molecular model of the SPKK peptide interacting with the minor groove of the DNA fragment.

The dashed lines represent the hydrogen bonds. The sugar-phosphate backbone of DNA is indicated. The peptide is shown in thick solid lines.

Table 4. The arrangement of hydrogen bonds for the homeodomain and SPKK peptide complexed with DNA

Donor	Acceptor	Distance (Å)
T8 (B) O2	RPRT Arg2 NH1	2.78
T8 (B) O2	RPRT Arg2 NH2	2.86
C9 (B) O1P	RPRT Arg2 NH2	2.62
C9 (B) O2P	RPRT Lys1 NZ	2.61
G10 (B) O1P	RPRT Lys1 N	2.56
G10 (B) O1P	RPRT Arg2 N	2.80
G4 (A) O2	RPRT Arg4 NH1	2.81
A6 (A) O1P	RPRT Thr5 OG1	2.40
T8 (A) O1P	RPRT Lys1 NZ	2.66
C9 (B) O1P	SPKK Lys1 NZ	2.85
C9 (B) O2P	SPKK Lys1 NZ	2.68
G10 (B) O1P	SPKK Lys1 N	2.53
G12 (B) O1P	SPKK Lys6 NZ	2.56
A5 (A) O4'	SPKK Lys4 NZ	2.89
A6 (A) O1P	SPKK Lys5 NZ	2.56
T8 (A) O1P	SPKK Lys1 NZ	2.64

The calculations of the electrostatic contribution to the free energy of binding for all generated models are summarized in Tables 5 and 6. In all cases the binding energies are negative, indicative of the stabilization of the analyzed binding modes. It is worth noting that the charge-charge Coulombic interaction energies and the solvent reaction field contribute in an opposite way to the total free energy of binding. Their absolute values are large and cancel each other to a large extent. The presence of counterions destabilizes the total binding energy. The destabilization is, however, smaller at equilibrium distances that at larger distances, therefore the ion cloud creates the binding effect at the van der Waals distances.

Table 5. The electrostatic free energy of binding for the model structures

Sequence	ΔG (kcal/mol)
Linear Poisson-Boltzmann model	
APEK	-2.135
APEK_H*	-7.836
SPEK	-1.814
SPEK_H	-6.119
Nonlinear Poisson-Boltzmann model	
APEK	-1.811
APEK_H	-6.027
SPEK	-1.227
SPEK_H	-4.415

*_H – protonated form of glutamic acid.

Table 6. The binding energy for the homeodomain fragment and canonical SPKK peptides computed using the nonlinear Poisson-Boltzmann equation

Sequence	ΔG (kcal/mol)
APKK//AT*	-7.153
APKK//CG**	-6.804
RPRT//AT	-7.947
RPRT//CG	-8.304
SPKK/AT	-4.774
SPKK//CG	-4.635

*//AT, DNA sequence: ATGTAATTACCT (from the crystallographic structure of homeodomain in complex with DNA); **//CG, DNA sequence: CGCAATTCGCG (the sequence of DNA used in the experiment).

The calculations performed for peptides with alanine in the location of serine at the position *i–1* show, that this replacement does

not affect the binding of the motif to the DNA sequences (Table 5 and 6). Interestingly, the free energy of binding in the case of the APEK peptide is lower (better binding) than in the case of the SPKK peptide. The finding that the Ser → Ala mutation makes the binding stronger suggests that the presence of serine is not a necessary condition for creating the specific interactions between DNA and the SPKK motif. The binding of the naturally occurring SPEK and APEK peptides that contain the deprotonated (negatively charged) glutamic acid is weaker than the binding of the peptides in which this residue is protonated (neutral). Upon the protonation of glutamic acid the stability of the complexes with the same peptides is considerably increased (Table 5).

DISCUSSION

The peptide motifs of the SPXZ-type, with X being any residue and Z a basic one, are the elements of many AT-rich DNA binding proteins, which bind to the minor groove. They are particularly abundant in histones of the H1 family in which the serine in the motif is considered a target of specific phosphorylation by mitotic histone H1 kinase.

In this work we analyzed the binding properties of the SPEK peptide to AT-rich DNA fragments. The peptide is a fragment of the N-terminal tail of the *A. thaliana* histone H1. This peptide competes with Hoechst 33258 for binding to the single site present in the specific DNA dodecamer and also binds efficiently to longer DNA with multiple AT-rich binding sites. In order to estimate the role for the above binding properties of serine at the *i−1* position we analyzed the APEK peptide that differs from the SPEK one by having alanine at the position of serine. The replacement does not change the binding properties of the motif. Thus, the presence of the conserved serine in this motif is probably not dictated by structural requirements. In order to interpret the experimental results we applied molecular modeling methods. First, the classical model for the SPXZ-peptides interacting with the AT-rich DNA was based on the assumption that the binding mode is essentially similar to that revealed for the netropsin–DNA complex [3]. However, in the course of the modeling, and in agreement with the earlier data [25], we concluded that the netropsin-type model is improbable and we proposed the binding mode based on the homeodomain N-terminal arm interaction with the AT-rich DNA. By screening the protein data base we found that the canonical SPKK motif fits well into the homeodomain N-terminal arm. Assuming the structural similarities among homeodomains of different proteins we chose the known crystallographic structure of the homeodomain of the engrailed protein in complex with DNA. The RPRT-peptide of the homeodomain N-terminal arm, that binds to the AT-rich site in DNA, was shown by us to bind to the DNA containing multiple AT-rich binding sites (Fig. 2). Based on the three dimensional structure of the complex between the RPRT peptide and DNA, we modeled the structures of the canonical (SPKK) and *A. thaliana* histone H1 specific (SPEK) peptides, as well as corresponding peptides containing the serine→alanine substitution, interacting with DNA. These structures did not produce any steric clashes between the peptides and DNA. In addition, the calculated electrostatic free energies of binding indicated the stabilization of all derived complexes. Similarly as in the DNA binding experiments, the replacement of serine by alanine did not affect the binding properties of the studied peptides. While the presence of non-protonated glutamic acid at the *(i + 1)* position led to a considerable decrease of the binding of peptides to AT-rich DNA, the results of the calculations for peptides with protonated form of glutamic acid showed a marked increase in the stability of the complex. Our model indicates that glutamic acid is probably protonated in the SPEK motif of the

DNA-complexed histone H1. Both the experiments and the modeling studies confirm that serine itself in the SPXZ motifs is not essential for the DNA binding. This is consistent with the idea that the conservation of serine in this motif is due entirely to the requirements of the phosphorylation process.

Note

The geometries of the modeled complexes can be obtained by e-mail (grycuk@icm.edu.pl).

REFERENCES

1. Pabo, C.O. & Sauer, R.T. (1984) Protein-DNA recognition. *Annu. Rev. Biochem.* **53**, 293-321.
2. Steitz, T.A. (1990) Structural studies of protein-nucleic acids interactions; The sources of sequence specific binding. *Quart. Rev. Biophys.* **25**, 205-280.
3. Suzuki, M. (1989) SPKK, a new nucleic acids binding unit of proteins found in histones. *EMBO J.* **8**, 797-804.
4. Reevs, R. & Nissen, M.S. (1990) The AT-DNA-binding domain of mammalian high mobility group chromosomal proteins. *J. Biol. Chem.* **265**, 8573-8582.
5. Strauss, F. & Varshavsky, A. (1984) A protein binds to a satellite DNA repeat at three specific sites that would be brought into proximity by DNA folding in nucleosome. *Cell* **37**, 889-901.
6. Kopka, M.L., Pjura, P.P., Yoon, C., Goosel, D. & Dickerson, R.E. (1985) Binding of antitumor drug to DNA. *J. Mol. Biol.* **183**, 553-563.
7. Coll, M., Frederick, C.A., Wang, . H.J. & Rich, A. (1987) A bifurcated hydrogen-bonded conformation in the d(AT) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8385-8389.
8. Omichinski, J.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S.J. & Gronenborn, A.M. (1993) NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* **261**, 438-446.
9. Kissinger, C.R., Liu, B., Martin-Blanko, E., Kornberg, T.B. & Pabo, C. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: A framework for understanding homeodomain-DNA interactions. *Cell*, **63**, 579-590.
10. Suzuki, M. (1989) SPXX, a frequent sequence motif in gene regulatory proteins. *J. Mol. Biol.* **207**, 61-84.
11. Shetty, K.T., Link, W.T. & Pant, C.H. (1993) cdc2-Like kinase from rat spinal cord specifically phosphorylates KSPXK motifs in neurofilamental proteins: Isolation and characterization. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6844-6848.
12. Tomaszewski, R. & Jerzmanowski, A. (1997) The AT-rich flanks of the oocyte-type 5S RNA gene of *Xenopus laevis* act as a strong local signal for histone H1-mediated chromatin reorganization *in vitro*. *Nucleic Acids Res.* **25**, 458-465.
13. Kandolf, M. (1994) The H1a histone variant is the *in vivo* repressor of oocyte-type 5S gene transcription in *Xenopus laevis* embryos. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7257-7261.
14. Shen, X. & Gorowski, M. (1996) Linker histones are not essential and affect chromatin condensation *in vivo*. *Cell* **86**, 475-483
15. Prymakowska-Bosak, M., Przewłocka, M.R., Iwkiewicz, J., Egierszdorff, S., Kuraś, M., Chaubet, N., Gigot, C., Spiker, S. & Jerzmanowski, A. (1996) Histone H1 overexpressed to high level in tobacco affects certain developmental programs but has limited effect on basal cellular functions. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10250-10255.

16. Russel, S.T. & Waishel, A. (1985) Calculations of electrostatic energies in proteins. The energetics of ionized groups in bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **185**, 389–404.

17. Misra, V.K., Sharp, K.A., Friedman, R.A. & Monig, B. (1994) Salt effects on ligand-DNA binding. Minor groove binding antibiotics. *J. Mol. Biol.* **238**, 245–263.

18. Zacharias, M., Luty, B.A., Davis, M.E. & McCammon, J.A. (1992) Poisson-Boltzman analysis of the λ repressor-operator interaction. *Biophys. J.* **63**, 1280–1285.

19. Misicka, A., Lipkowski, A.W., Barcikowska, M. & Pluta, R. (1995) Synthesis and formulation for biological study of β amyloid (1–42) polypeptide; in: *Peptides 1994*, Proceedings of the 23rd European Peptide Symposium (Maia, H.L.S., ed.) pp. 377–378, ESCOM, Leiden.

20. Lesyng, B. & McCammon, J.A. (1993) Molecular modeling methods. Basic techniques and challenging problems. *Pharmac. Ther.* **60**, 149–167.

21. McCammon, J.A. & Harvey, S.C. (1987) *Dynamics of Proteins and Nucleic Acids.* Cambridge University Press, Cambridge.

22. Grycuk, T. & Lesyng, B. (1996) Studies on macromolecular recognition processes. Application of a Poisson-Boltzman model to protein-nucleic acids interactions. *3rd European Conference on Mathematics Applied to Biology and Medicine*, October 6–10, Abstract, p. 29, Heidelberg.

23. Churchill, M.E.A. & Suzuki, M. (1989) 'SPKK' motifs prefer to bind to DNA at A/T-rich sites. *EMBO J.* **8**, 4189–4195.

24. Wilkinson, G.N. (1961) Statistical estimations in enzyme kinetics. *Biochem J.* **80**, 324–332.

25. Suzuki, M., Gerstein, M. & Johnson, T. (1993) An NMR study on the DNA-binding SPKK motif and a model for its interaction with DNA. *Protein Eng.* **6**, 565–574.