BIOGERONTOLOGY

Study of Interactions between DNA and Tetrapeptides using Methods of Molecular Mechanics

S. I. Tarnovskaya[1], P. P. Yakutseni[2], and V. Kh. Khavinson[1,3]

Translated from *Byulleten' Eksperimental'noi Biologii i Meditsiny*, Vol. 156, No. 11, pp. 637-641, November, 2013
Original article submitted November 13, 2012

Tetrapeptide Lys-Glu-Asp-Trp is a mimetic of insulinotropic peptides and reduces the blood glucose level. This tetrapeptide increases the content of important factors of differentiation in endocrine pancreatic cells *in vitro*. Molecular modeling shows that this tetrapeptide can interact with not only minor, but also major groove of DNA molecule. The interaction with the major groove is more specific, because it depends on the primary sequences of the tetrapeptide and DNA. Sequence GGCAG is the putative binding site for the tetrapeptide.

Key Words: *tetrapeptide Lys-Glu-Asp-Trp; DNA duplex; diabetes mellitus; docking; molecular mechanics*

Physiological activity of pancreatic endocrine cells decreases with age, which leads to the development of hyperglycemia progressing to diabetes mellitus. Dysfunction of insulin-producing β-cells is one of the causes of diabetes. The decrease in physiological activity of β-cells during aging can be due to impaired synthesis of specific proteins involved in the regulation of cell proliferation and differentiation.

Tetrapeptide Lys-Glu-Asp-Trp exhibits an insulinotropic effect [1,2,8,15]. Experiments on rats with alloxan-induced diabetes mellitus showed that this tetrapeptide reduced blood glucose level. In cultured human pancreatic cells, tetrapeptide promoted synthesis of proteins Pdx1, Pax6, Nkx2.2 and Ngn3, important factors of β-cells differentiation [1].

Most of the known peptides (GLP1, bZIP, CPP, Lys-Trp-Lys-Lys, *etc.*) exhibit a wide spectrum of biological effects mediated through binding to DNA molecules [4,7,10,11,13]. Peptides interacting with DNA molecule induce its conformational changes resulting in activation or inhibition of gene expression. Under physiological conditions, these peptides are protonated and positive charge promotes their binding to negatively charged sites of the minor groove of DNA [3,7]. However, such interactions are not enough specific and almost do not depend on nucleotide sequence in DNA, which attests to common mechanism of their action.

Previous data suggest that tetrapeptide Lys-Glu-Asp-Trp activated the expression of genes encoding β-cell differentiation factors by interacting with a specific binding site in the major groove of DNA molecule [1]. Sequence GGCAG in the promoter region of genes encoding proteins Pdx1, Pax6, Pax4, Nkx2.2, and Ngn3 and in preproinsulin is a putative binding site for tetrapeptide Lys-Glu-Asp-Trp.

Addition of another tetrapeptide, Ala-Glu-Asp-Leu, with amino acid substitutions in the first and fourth positions to the culture of pancreatic cells did not increase the synthesis of differentiation factors of β-cells. This can indicate that the peptide does not

[1]St. Petersburg Institute of Bioregulation and Gerontology; [2]Center for Advanced Studies, St. Petersburg State Polytechnic University; [3]I. P. Pavlov Institute of Physiology, Russian Academy of Sciences, St. Petersburg, Russia. ***Address for correspondence:*** svetlanatarnovskaya@gmail.com. S. I. Tarnovskaya

interact with specific binding sites and does not participate in activation of gene expression.

Computer simulation of DNA-peptide complexes based on our previous data and published reports can not only substantiate the selection of the optimal picture of DNA-peptide complex formation, but also serve for primary analysis of possible mechanisms of action of newly synthesized compounds.

The aim of the work was to develop computer models simulating the interactions between tetrapeptides Lys-Glu-Asp-Trp and Ala-Glu-Asp-Leu and DNA duplex and to reveal differences in their binding.

MATERIALS AND METHODS

The Molecular Operating Environment 2012 software (MOE 2012.10) was used for molecular modeling [12].

Optimal dimensional structures of tetrapeptides Lys-Glu-Asp-Trp and Ala-Glu-Asp-Leu were examined by conformational search with molecular dynamics simulations using MM, CHARMM27, Amber99 and other force fields. The main results were obtained using MMFF94 force field [6]. The calculations were performed using standard schemes of molecular mechanics considering (or not considered) the solvent. The effect of the solvent was calculated using generalized Born approximation with introduction of the internal dielectric constant equal to 1 and external constant equal to 80. The first 56 of 200 obtained peptide conformers were selected, which corresponded to studied range of 7 kcal/mol. Further, protein-DNA docking was performed using these conformers. As a double-stranded DNA molecule, B-form DNA duplex 5'-ATATGGCAGGGGTAA-3' containing a putative binding site for GGCAG was chosen. To simulate a complex of double-stranded DNA with tetrapeptides Lys-Glu-Asp-Trp and Ala-Glu-Asp-Leu, molecular docking method was used consisting in selection of the optimal mutual orientation of the molecules in their binding and the formation of a stable complex. When predicting optimal peptide orientation in DNA duplex, contact area, number of hydrogen bonds, parameters of hydrophobic and electrostatic interactions were taken into account. We used the standard search algorithm in MOE 2012.10, the whole DNA molecule was chosen as a binding site.

Geometry optimization of the complex was performed for every docking solution. Binding energy calculated as the difference between the potential energies of DNA-tetrapeptide complex and single molecules. Next, the entropy contribution (12 kcal/mol) was subtracted. The intensity of interaction of the peptide with DNA was determined by the magnitude of DE negative values: the interaction is stronger at lower negative energy.

RESULTS

We performed docking of tetrapeptides Lys-Glu-Asp-Trp and Ala-Glu-Asp-Leu in DNA duplex:
5'-$A^1T^2A^3T^4G^5G^6C^7A^8G^9G^{10}G^{11}G^{12}T^{13}A^{14}A^{15}$-3'
3'-$T^{30}A^{29}T^{28}A^{27}C^{26}C^{25}G^{24}T^{23}C^{22}C^{21}C^{20}C^{19}A^{18}T^{17}T^{16}$-5'

The main docking results were obtained in MMFF94 force field. We present results for the complexes with binding energy <-5 kcal/mol (Table 1). Complexes with binding energy >-5 kcal/mol are unlikely. Tetraprptide Lys-Glu-Asp-Trp interacts with both the minor and major grooves of DNA molecule. The interaction of the tetrapeptide Lys-Glu-Asp-Trp with the minor groove of DNA proved to be energetically most favorable: binding energy of this complex was -14.2 kcal/mol. In the minor groove, the tetrapeptide formed hydrogen and ionic bonds between amino groups of Lys and oxygen atoms of phosphate groups and deoxyribose C^{21} and G^{12}. DNA-peptide complex located in the major groove of DNA was the second by the value of binding energy. The peptide formed two specific hydrogen bonds between N7 G^5 and the terminal amide group of the peptide as well as between the carbonyl group of the main chain of Asp peptide and N4 C^{25}. Hydrogen bonds between the two NH_3^+ groups of Lys and oxygens in phosphate residues T^{23} and G^{24} also were formed. The binding energy of this complex was -13.71 kcal/mol. In the major groove of the DNA, the peptide binds to side groups of DNA nitrogenous bases forming specific contacts. Thus, NH of Trp side chain non-covalently interacted with guanine O6 atom and adenine N7 atom, carboxyl groups of Asp and Glu interacted with cysteine N4 and guanine (Fig. 1, *a-c*). Lys side groups interacted with nucleotides of the minor groove of DNA forming hydrogen and ionic bonds with the oxygen atoms of the phosphate groups and deoxyribose in adjacent nucleotides.

Specific contacts formed between $NH3^+$ of Lys side group and O2 T^4 and N3 A^{27} also occurred, though less frequently than in the peptide interaction with the major groove of DNA (Fig. 2, *a*).

Thus, 9 binding sites for peptide Lys-Glu-Asp-Trp to DNA are included in sequence GGCAG in the putative specific binding site for tetrapeptide Lys-Glu-Asp-Trp, and 6 of them are located in the major groove of DNA.

We performed docking of tetrapeptide Ala-Glu-Asp-Leu, which has no biological effect on pancreatic cells, to DNA duplex. Twenty docking solutions were generated, 18 of them were located in the minor groove, and 2 in the major groove. First 12 docking solutions are presented and the lack of energetically favorable interactions was shown (Table 2). The maximum interaction energy for the peptide-DNA complex was -3.18 kcal/mol, which is insufficient to form a

TABLE 1. Results of Docking of Tetrapeptide Lys-Gly-Asp-Trp to DNA Molecule with Binding Energy Values below -5 kcal/mol

Duplex DNA 5'→3'															ΔE, kcal/mol	DNA groove
A	T	A	T	G	G	C	A	G	G	G	G	T	A	A		
							*	*	*						-14.2	minor
				*	*	*									-13.71	major
									*	*	*	*			-12.25	minor
		*	*	*	*										-8.82	minor
*	*	*	*												-8.55	minor
					*	*	*	*							-8.13	major
					*	*	*								-6.4	major
			*	*	*										-6	major
					*	*	*								-5.98	major
		*	*	*											-5.48	minor
					*	*	*								-5.32	major
								*							-5	minor

Note. Here and in Table. 2: asterisks mark the sites of interaction between tetrapaptide and DNA. ΔE: free energy.

TABLE 2. Results of Docking of Tetrapeptide Ala-Gly-Asp-Leu to DNA

Duplex DNA 5'→3'															ΔE, kcal/mol	DNA groove
A	T	A	T	G	G	C	A	G	G	G	G	T	A	A		
	*	*	*												-3.57	minor
	*	*	*												-3.18	minor
	*	*													-1.92	minor
	*	*	*												-1.68	minor
	*	*	*												-1.55	minor
	*	*													1.35	minor
	*	*	*												-1.12	minor
	*	*													-0.7	minor
*	*														-0.51	minor
	*	*													-0.24	minor
	*	*	*												0.33	minor
	*	*	*												1.38	minor

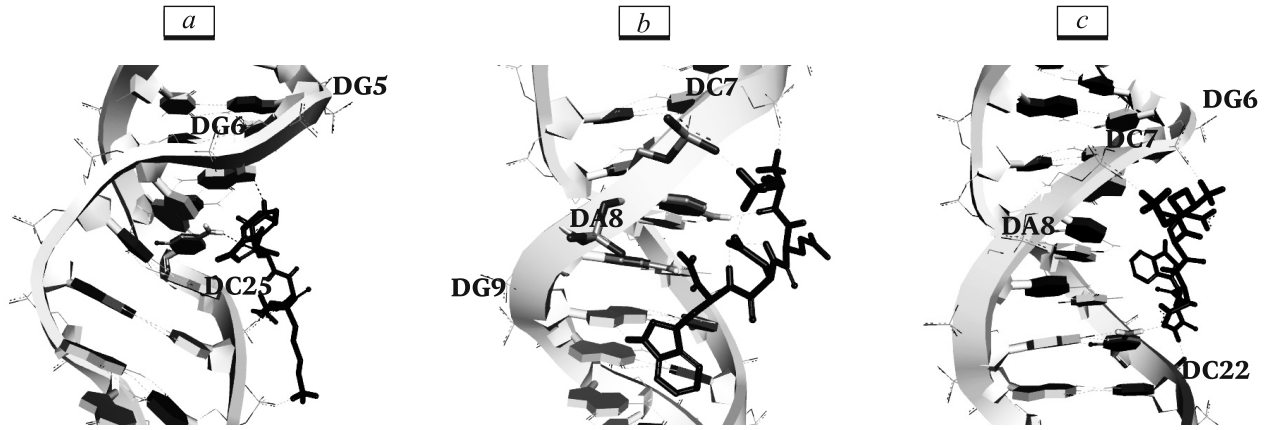

Fig. 1. Tetrapeptide Lys-Glu-Asp-Trp in the major groove of duplex DNA formed by the sequence GGCAG. *a*) Complex with ΔE=-13.7 kcal/mol, *b*) complex with ΔE=-8.13 kcal/mol, *c*) complex with ΔE =-5.9 kcal/mol. Dotted line, hydrogen bonds between molecules of peptide and DNA.

stable complex. Tetrapeptide Ala-Glu-Asp-Leu contacted with other DNA sequence, 5'TAT3', which does not match the binding site for Lys-Glu-Asp-Trp (Table 2; Fig. 1, *b*).

Molecular modeling explored the possibility of interaction between peptides (Lys-Glu-Asp-Trp and Ala-Glu-Asp-Leu) and double helix of DNA. Under physiological conditions, peptides have a negative charge due to deprotonation of glutamic and aspartic acids and their specific interaction with the polyanion (DNA molecule) seems unlikely. However, docking considering electrostatics showed options for peptide interactions both in the major and the minor groove of DNA. We revealed binding specificity of the tetrapeptide Lys-Glu-Asp-Trp with sequence GGCAG, and tetrapeptide Ala-Glu-Asp-Leu, with sequence TAT.

Frequency analysis of the amino acid residues interacting with the nucleotide bases in the three-dimensional protein complexes with polynucleotides indicated that the frequency of negatively charged glutamic and aspartic acids was very low and differs from that of the arginine and lysine [3,5,9,14]. However, in the structures both of investigated tetrapeptides both acid occur forming hydrogen bonds with the amino groups of cytosine and guanine.

All the complex of type "peptide in the minor groove of DNA" are characterized by a lack of specificity with respect to nitrogenous bases, and the formation of hydrogen bonds and electrostatic bonds is the most distinctive type of interaction. On the contrary, interactions between peptides and major groove of DNA is specific because includes not only the formation of hydrogen bonding by side groups of peptide residues with DNA nitrogen bases, but also a wider set of intermolecular interactions (electrostatic and hydrogen bonding, hydrophobicity, more pronounced van der Waals forces and others.). Moreover, such interactions were found only in complex Lys-Glu-Asp-Trp with DNA, which indicates the selective effect of above tetrapeptide and defines its distinctive features among most other previously studied peptides.

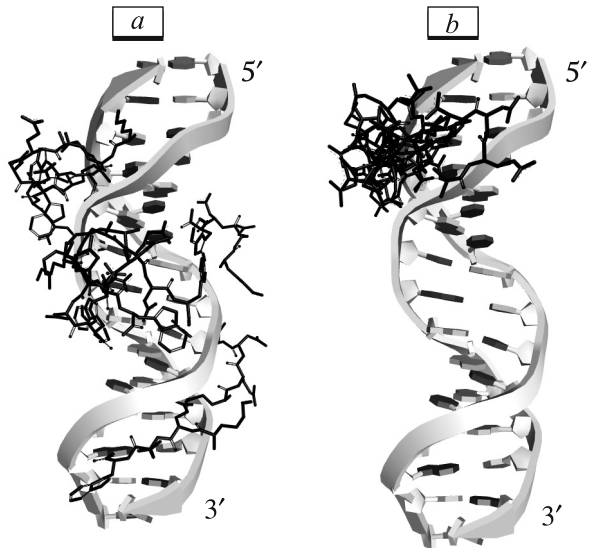

Fig. 2. Set of tetrapeptide docking solutions in 15 b.p. DNA duplex. *a*) Complexes of Lys-Glu-Asp-Trp with DNA; *b*) complexes of Ala-Glu-Asp-Leu with DNA.

REFERENCES

1. V. Kh. Khavinson, A. O. Durnova, V. O. Polyakova, *et al*., *Bull. Exp. Biol. Med*., **154,** No. 4, 501-504 (2012).
2. V. N. Anisimov and V. Kh. Khavinson, *Biogerontology*, **11**, No. 2, 139-149 (2010).
3. A. C. Cheng, W. W. Chen, C. N. Fuhrmann, and A. D. Frankel, *J. Mol. Biol.,* **327**, No. 4, 781-796 (2003).
4. A. Chung, F. Eudes, and Y.-S. Shim, *IUBMB Life*, **62**, No. 3, 183-193 (2010).
5. B. H. Geierstanger, M. Mrksich, P. B. Dervan, and D. E. Wemmer, *Science*, **266**, 646-650 (1994).
6. N. A. Halgren and R. B. Nachbar, *J. Comput Chem*., **17**, 587-615 (1996).

7. H. Huang, I. D. Kozekov, A. Kozekova, *et al*., *Biochemistry*, **49**, No. 29, 6155-6164 (2010).
8. V. Kh. Khavinson and V. V. Malinin, *Gerontological Aspects of Genome Peptide Regulation*, Basel (2005).
9. V. Kh. Khavinson, A. Yu.Solov'ev, D. V. Zhilinskii, *et al*., *Adv. Gerontol*., **2**, No. 4, 11-22 (2012).
10. N. M. Luscombe, R. A. Laskowski, and J. M. Thornton, *Nucleic Acids Res*., **29**, No. 13, 2860-2874 (2001).
11. P. E. MacDonald, W. El-kholy, M. J. Riedel, *et al*., *Diabetes*, **51**, Suppl. 3, S434-S442 (2002).
12. *Molecular Operating Environment*; *Chemical Computing Group Inc* (2012) 1010 Sherbooke St. West, Suite No. 910, Montreal, QC, Canada, H3A 2R7.
13. S. Ryu, S. Zhou, A. G. Ladurner, and R. Tjian, *Nature*, **397**, 446-450 (1999).
14. S. Trabulo, A. L. Cardoso, M. Mano, and M. C. PedrosoDelima, *Pharmaceuticals*, **3**, No. 4, 961-993 (2010).
15. V. Kh. Khavinson, V. V. Malinin, E. I. Grigoriev, and G. A. Ryzhak, *USA Patent* US 7,491,703. *Tetrapeptide regulating blood glucose level in diabetes mellitus*, (17.02.2009).