Stick of Sticks: Structural Features of the Amyloidogenic Peptide-DNA Complex

Arzamastsev G.A.[1], Zabrodskaya Y.A.[1,2], Garmay Yu.P.[3], Ivanova N.S.[4], Vinogradova D.S.[3], Arutyunyan A.V.[3], Verlov N.A.[3], Burdakov V.S.[3], Baymukhametov T.N.[5], Konevega A.L.[1,3,5], Ivankov O.I.[6], Gorshkova Yu. E.[6], Egorov V.V.[1,4]

[1] Institute of Biomedical Systems and Biotechnology, Peter the Great St. Petersburg Polytechnic University, 194064, Politekhnicheskaya 29, St. Petersburg, Russia

[2] Smorodintsev Research Institute of Influenza, Russian Ministry of Health, 197376, Prof. Popov St. 15/17, St. Petersburg, Russia

[3] Petersburg Nuclear Physics Institute named by B.P. Konstantinov of National Research Centre "Kurchatov Institute", 188300, mkr. Orlova roshcha 1, Gatchina, Russia

[4] Federal State Budgetary Scientific Institution 'Institute of Experimental Medicine', 197376, Akademika Pavlova street 12, St. Petersburg, Russia

[5] National Research Centre Kurchatov Institute, 123182, Akademika Kurchatova Sq. 1, Moscow, Russia

[6] International Intergovernmental Organization Joint Institute for Nuclear Research, 141980, Joliot-Curie St. 6, Dubna, Moscow region, Russia

Corresponding author:

Vladimir Egorov, sondyn@yandex.ru

Abstract

The search for peptides that can specifically bind to regulatory regions in DNA is a necessary step for creating drugs that can regulate gene expression. The work studies the peculiarities of binding of a model peptide, which carries an ionic self-complementary motif and is capable of forming amyloid-like fibrils, with model double-stranded DNA. The stoichiometric ratios of the components of the complex were found using the retardation method in agarose gel. Using microscale thermophoresis, it was shown that the peptide in the oligomeric amyloid-like state is capable of binding to model 45-bp double-stranded DNA, with a micromolar equilibrium dissociation constant. Using electron and atomic force microscopy, the morphology of peptide-DNA complexes was studied; using dynamic light scattering and nanoparticle tracking analysis, as well as small-angle neutron scattering, the spatial parameters of the resulting DNA-peptide complexes were characterized. Molecular dynamics simulations showed that the arginine side chains of the peptide are prone to interact with guanine nitrogenous bases. It was shown that the formation of peptide-dsDNA complexes interferes with the operation of restriction endonucleases that have guanine-cytosine pairs in the recognition center, which is consistent with the results of prediction of interaction sites obtained using computer modeling. The results of the work can be used in the development of peptides capable of interacting with functional regions of DNA, as well as in the development of new carriers for transfection of DNA constructs.

1. Introduction

The design of drugs based on substances capable of regulating gene expression is an urgent task [1]. A number of structures of natural transcription regulators are known [2], [3], both used by host cells and viral ones, which allow influencing gene expression during viral infection [4], [5], [6]. Small synthetic peptides are also known that can specifically interact with certain DNA structural elements [7], [8]. Proteins and peptides capable of amyloid-like oligomerization necessary to perform their functions are called functional amyloids [9], [10]. The prion-like properties of such polypeptides are also realized in processes where positive feedback is necessary - the reaction of a living system to small amounts of stimulus [11], [12]. Amyloid-like fibrils of a number of proteins are capable of interacting with DNA; DNA can also initiate conformational transitions necessary for the protein to acquire a beta conformation and oligomerization [13]. The formation of functional amyloids also occurs in some cases of epigenetic regulation. Recently, it was shown that similar mechanisms of expression regulation are realized both in yeast [14], and in mammals [15]. Of particular interest are supramolecular peptide complexes having a regular structure, especially, peptides capable of forming amyloid-like fibrils and of cooperative interaction with DNA [16]. Ionic self-complementary motifs are elements of the primary structure of polypeptides, including periodically located amino acid residues with alternating positive and negative charges of side chains [17], [18], [19], [20]. Although such motifs are often found in the functional domains of proteins, including DNA-binding ones, they are poorly represented in databases of proteins that modulate gene expression. In addition, peptides bearing ionic self-complementary motifs have the potential to be used as therapeutic nucleic acid delivery vehicles. [21]. In this work, we study the structural prerequisites for the specific interaction of amyloid-like fibrils formed by a model peptide carrying an ionic self-complementary motif with model DNA sequences.

2. Materials and methods

2.1 Peptide IRIDIRIDIRG (SID)

Peptide with amino acid sequence IRIDIRIDIRG (called SID) was synthesized was synthesized by "Verta" Research and Production Company (Russia), the purity is more than 95%. Reagents used were produced by Sigma Aldrich (USA) or New England Biolabs, Inc (USA). Double-stranded DNA (dsDNA) (1P, 45 base pairs) was synthesized by DNA Synthesis (Russia), and labeled with Cy5 along the plus strand at the 3' end. The sequence of such DNA is given in Table 1. Lambda phage DNA was produced by SibEnzym (Russia).

Table 1. 1P DNA sequence

```
5′ -GGGTCGAGGGTCGAGGGTCGAGGGTCGAGGGTCGAGGGTCGAGGG- Cy5 - 3′
5′ -CCCTCGACCCTCGACCCTCGACCCTCGACCCTCGACCCTCGACCC- 3′
```

2.2 Preparation of fibrils from SID peptide

To obtain fibrils, the SID peptide was dissolved in 100% dimethylsulfoxide (DMSO; Dimexid, Tula Pharmaceutical Factory LLC, Russia), at a concentration of 10 mg/ml, then HEPES buffer (Sigma, USA) was added to a concentration of 1 mg/ml for the peptide; the mixture was incubated at 55°C for 180 min.

To select fibrillogenesis conditions, nanoparticle tracking analysis (NTA) were performed (see sec. 2.10) using both registration of particle trajectories in transmitted light and only fluorescent particles upon excitation at 488 nm. Complexes obtained under the conditions described above were registered identically under both measurement conditions, indicating the presence in the solution of only particles capable of binding Thioflavin T (ThT), i.e., of an amyloid-like nature. In further experiments, only fibrils obtained under such conditions were used.

2.3 Preparation of peptide-DNA complexes

The peptide-DNA complexes under study were obtained by mixing dsDNA dissolved in HEPES buffer at the appropriate concentration with SID peptide fibrils. The resulting mixtures were incubated at a temperature of 55 °C for 6 hours in a BioRad thermostat (USA).

2.4 Atomic force microscopy (AFM)

All reaction mixtures were diluted 10 times with a buffer for DNA binding (4 mM Tris-HCl (pH 7.5), 4 mM $NiCl_2$) [22]. 10 μL of the diluted sample was applied to freshly cleaved 2SPI mica (USA) and incubated for 3 minutes at room temperature. The samples were then washed with deionized water and dried in a gentle stream of compressed air. AFM images were obtained in air

using an atomic force microscope NT-MDT Solver BIO (Russia) with a Smena-B measuring head operating in semi-contact mode, using cantilevers NT-MDT NSG01, NSG03 (Russia), at a scanning frequency of 1 Hz and resolution 512x512 pixels. Surface topography of DNA and DNA-peptide complexes were analyzed using the NT-MDT software package. Images were processed in the Gwyddon software package [23].

2.5 Thioflavin T (ThT)

Amyloid-like fibrils were detected using ThT [24]. Thioflavin T (Sigma) was dissolved in phosphate buffered saline (PBS) and filtered through a 0.2 μm syringe filter. ThT concentration was then determined using an extinction coefficient of 36 mM^{-1} cm^{-1} at a wavelength of 412 nm. To measure ThT fluorescence, samples of fibrillar complexes of the peptide and its complexes with lambda phage DNA were prepared in PBS buffer. 2 mL of 10 μM ThT solution was mixed with the peptide, model DNA, and DNA-peptide complexes and transferred to a cuvette for fluorescence measurement. ThT fluorescence was measured at room temperature using a Hitachi7000 fluorimeter with an excitation wavelength of 440 nm and recording at 450–700 nm.

2.6 Congo red (CR)

Congo red dye produced by Sigma Aldrich (USA) was used. A 500 μM (10-fold) dye solution in PBS was prepared and stored. Before measurement, a 5-fold molar deficiency of the sample (10 μM) was added to the dye diluted 10 times with PBS [25]. The absorption spectrum was obtained on a Hitachi 3310 instrument using disposable cuvettes with a volume of 100 μl.

2.7 Agarose gel retardation

Peptide fibrils at the appropriate concentration in HEPES buffer were added to double-stranded model DNA and incubated for 180 minutes at 55°C. Then samples were mixed with gel loading dye (New England Biolabs) and was applied to the 1.5% agarose gel wells (in TE buffer). Electrophoresis was carried out at 130 mA for 1.5 h at a constant temperature of 20°C. Visualization of the bends was carried out using the BioRad ChemiDoc gel documentation system (BioRad, USA) sequentially in the fluorescence detection mode of ethidium bromide and Cy5.

2.8 Microscale thermophoresis (MST)

To conduct an experiment to determine the binding constant of peptide monomers to DNA, samples of DNA complexes (1P) with peptide fibrils (SID) were prepared in various concentrations. DNA was dissolved in a 20 μl HEPES buffer solution at a concentration of 50 μM. Next, a model peptide dissolved in 100% DMSO was added to each of the 16 samples in

concentrations from 0.021 nM to 700 nM; the DMSO concentration in all samples was 2% v/v. To form complexes, the resulting mixture was left for 180 minutes at a temperature of 55°C. The equilibrium dissociation constant KD was calculated from the obtained curve (using equation [1])

$$[1]\quad F = F_u + \frac{(F_b - F_u)}{2} \cdot (C_f + C + 4) - \sqrt{(C_f + C + K_d)^2 - 4C_f \cdot C}$$

MST experiments were carried out at the National Research Center "Kurchatov Institute" – PNPI on a Monolith NT.115 device (NanoTemper Technologies GmbH, Germany) in accordance with the manufacturer's protocols.

2.9 Dynamic light scattering (DLS)

Experiments to determine the hydrodynamic radius using DLS were carried out for DNA fragments (1P) 45 nucleotide pairs long in complex with the SID peptide. The formation of complexes occurred in standard HEPES buffer for 180 minutes at a temperature of 55°C. DLS measurements were carried out using a laser correlation spectrometer LKS03 (Intoks, Russia) equipped with a helium-neon laser with a wavelength of 633 nm. LKS03 software (Russia) was used for data collection and analysis. Measurements were carried out in disposable polystyrene cuvettes with an optical path length of 10 mm; the sample volume was 500 µl. Measurements were carried out at a distance of 4.65 mm from the cell wall with an automatic attenuator and at a controlled temperature of 25°C. For each sample, 5 measurements of 180 seconds were performed.

2.10 Nanoparticle tracking Analysis (NTA)

NTA measurements were performed on a NanoSight 2.3 (USA) equipped with a sample chamber with a 640 nm laser and a Viton fluoroelastomer O-ring. Samples were injected into the sample chamber using sterile syringes until the liquid reached the tip of the nozzle. All measurements were carried out at room temperature. NTA 2.0 Build 127 software was used for data acquisition and analysis. Samples were measured for 180 s in High Dynamic Range mode, which splits the captured video into two videos with independent shutter speed and gain settings. Scattered light modes (wavelength 640 nm) and excitation mode at a wavelength of 406 nm and registration of only fluorescent particles were used (using a buffer containing 10 µg/ml Thioflavin T). The average particle size and standard deviation values were obtained using NTA software.

2.11 Restriction analysis

Determination of the location of specific peptide binding sites on DNA was carried out using electrophoretic separation of DNA fragments obtained during enzymatic hydrolysis of DNA. To carry out the experiment, 3 enzymes were used: Ava1, Hind3, Nde1 in the corresponding

rCutSmart and 2.1 buffers (New England Biolabs) at concentrations recommended by the manufacturer and lambda phage DNA. DNA fragments were visualized by fluorescent DNA staining using ethidium bromide after electrophoretic separation in a 1.5% agarose gel.

2.12 Small angle neutron scattering (SANS)

SANS spectra were obtained on the YuMO spectrometer located on the fourth channel of the IBR-2 reactor (Frank Laboratory of Neutron Physics, Joint Institute for Nuclear Research, Dubna, Russia). To measure scattering spectra using the SANS method, solutions of peptides and peptide–DNA complexes were prepared in D_2O at a concentration of 1 mM. The measurements were carried out according to the standard method for this device, described earlier [25]. Scattering was recorded by two detectors simultaneously in the Q region from 0.006 to 0.3 Å.

Preliminary data processing was carried out according to the scheme described in [26]. Obtaining and normalizing the curves in absolute units was carried out using vanadium located directly in front of the detector. During the spectra analyses the following relationship should be fulfilled:

$$qI(q) \cong I_c(0)\exp\left(-\frac{1}{2}R_c^2q^2\right) \quad [2]$$

Based on this ratio, to determine the average radius of the resulting complexes, the spectra were plotted in Guinier coordinates for fibrillar structures (lnIq versus q^2, Figure 11), after which a linear approximation of the initial sections of the scattering curves was performed using the least squares method (using the Origin2018 software package) and the slope angle for each of the lines was calculated.

From formula [2] it follows that the radius Rc can be calculated:

$$\mathrm{tg}(\alpha) = -\frac{1}{2}\cdot R_c^2 \Rightarrow R_c = \sqrt{-2\cdot tg(\alpha)} \quad [3]$$

The Origin2018 software package was used to visualize the curves.

2.13 Molecular dynamics simulation

The results of computer modeling were obtained using the computing resources of the supercomputer center of Peter the Great St. Petersburg Polytechnic University (www.spbstu.ru).

As the initial state, we used fully unfolded conformations of the IRIDIRIDIRG peptide and unfolded 1P DNA in the B-form, generated in the PyMol program [27]. Using GROMACS [28] a system was built in the form of a cubic box with dimensions of 100^3 $Å^3$, containing 90 monomers of the SID peptide and Na^+ and Cl^- atoms in a number corresponding to a 25 mM NaCl

concentration. For peptides the field amber14sb was used [29], and tip3p for water [30]. The prepared starting configuration of the polypeptide was placed in a box with water molecules.

Next, the stages of minimizing the energy of the system and the first stage of balancing were carried out. Initial velocities were calculated based on the Maxwell distribution for temperature 310 K. The Berendsen thermostat and barostat were used in the calculations [31]. Solvent equilibration was performed during a 5 ns simulation with immobile protein. The last obtained configuration was used as a starting one at the second stage of equilibration, and the effect of potentials limiting the movement of atoms was removed. Used a Nose-Hoover thermostat [32], [33], [34], [35] and Parinello-Raman barostat [36], [37]. Equilibration was carried out for 10 ns. Next, molecular dynamics simulation was performed for 150 ns. The results of this modeling stage were used for further analysis. The obtained data was processed using the numpy and matplotlib packages. Images of the simulated trajectory were processed in the PyMol software package [38].

2.14 Cryo-electron microscopy (cryo-EM)

For cryo-EM sample preparation a 3 µl of protein solution was applied to the Quantifoil R1.2/1.3 300 mesh Cu grid, which was glow-discharged for 30 s at 15 mA using PELCO easiGlow (Ted Pella, USA). The grid was plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific, USA) with the following settings: chamber humidity 100 %; chamber temperature 4 °C; blotting time 3 s; blotting force 0. Cryo-EM data were collected on a Titan Krios 60-300 transmission electron microscope (Thermo Fisher Scientific, USA) with a Falcon II direct electron detector (Thermo Fisher Scientific, USA) at the National Research Centre “Kurchatov Institute”. The microscope was operated at 300 kV with a nominal magnification of 75000x, corresponding to a callibrated pixel size of 0.86 Å, total exposure dose was 80 $e^{-}/Å^{2}$ and fractionated into 32 frames during 1.6 s exposure time. A total of 400 micrographs were collected and pre-processed using Warp software [39]. Filament segments were picked using CrYOLO [40] and exported to the CryoSPARC [41]for 2D classification and helical reconstruction, resulting in a low-resolution cryo-EM map of SID peptide complexes.

3. Results and discussion

3.1 Peptide SID (IRIDIRIDIRG) is capable of forming amyloid-like fibrils

Previously we showed [20], that the GDIRIDIRIDIRG peptide is capable of forming amyloid-like fibrils. We also showed that GDIRIDIRIDIRG fibrils are capable of dissociation into oligomers in the presence of a guanine analogue, the antiviral drug Triazavirin [42]. Data obtained from molecular dynamics simulations in free diffusion mode indicated that the side chains of arginines are prone to interact with triazavirin heterocycles. We hypothesized that the interaction occurs through hydrogen bonding between the amines of arginine and triazavirine, as is the case with guanine [43]. Although guanine atoms in double-stranded DNA are involved in hydrogen bonding with cytosine, we hypothesized that perhaps the arginines in the peptide would be able to interact with guanine heterocycles.

The peptide GDIRIDIRIDIRG, which we previously studied, had low solubility, apparently due to electrical neutrality at physiological pH values. Therefore, despite the unique properties of such a peptide, the prospects for its use remained vague. In this work, we study a shortened version of the peptide GDIRIDIRIDIRG – IRIDIRIDIRG (called SID), which retains self-complementarity, but has a non-zero total charge at neutral pH and is more soluble. Figure 1 shows the results of atomic force (Fig. 1A) and cryo-electron (Fig. 1B) microscopy. In Figure 2 there are results of fluorimetry in the presence of the thioflavin T dye characteristic of amyloid-like fibrils (Fig. 2A), spectrophotometry in the presence of Congo red (Fig. 2B). The morphology of peptide aggregates, the ability to form fluorescent complexes with thioflavin T, as well as the ability of peptide aggregates to shift the absorption spectrum of Congo red to the right indicate that the truncated peptide IRIDIRIDIRG, just like the previously studied GDIRIDIRIDIRG, is capable of forming amyloid-like fibrils.

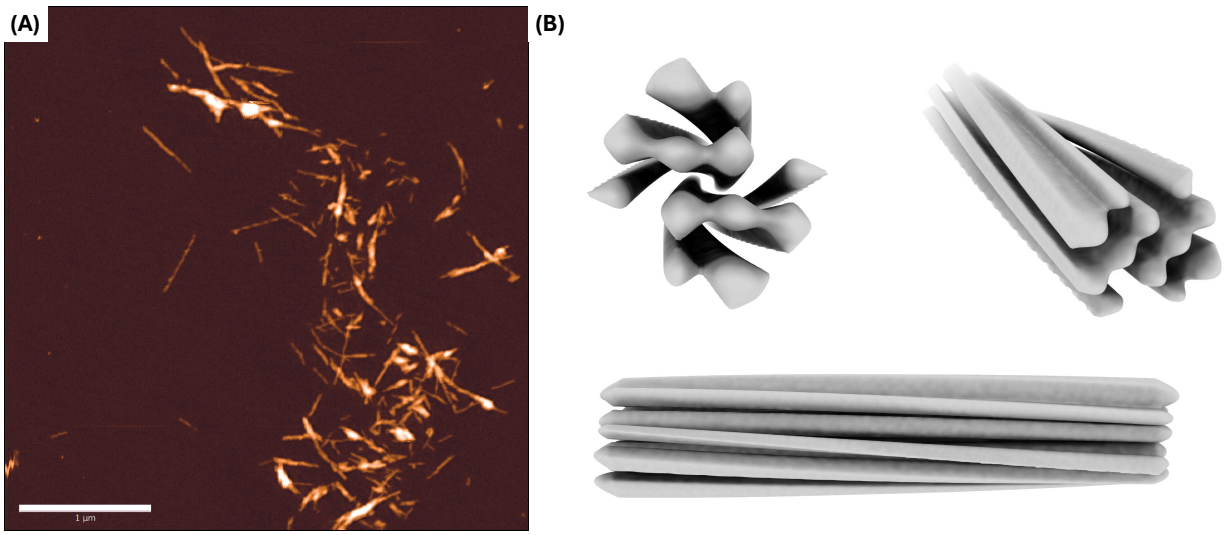

Figure 1. Microscopy of fibrillar structures formed by SID peptide. (A) Surface topography obtained using atomic force microscopy (scale bar is 1 um) (B) Reconstruction obtained from cryo-electron microscopy data.

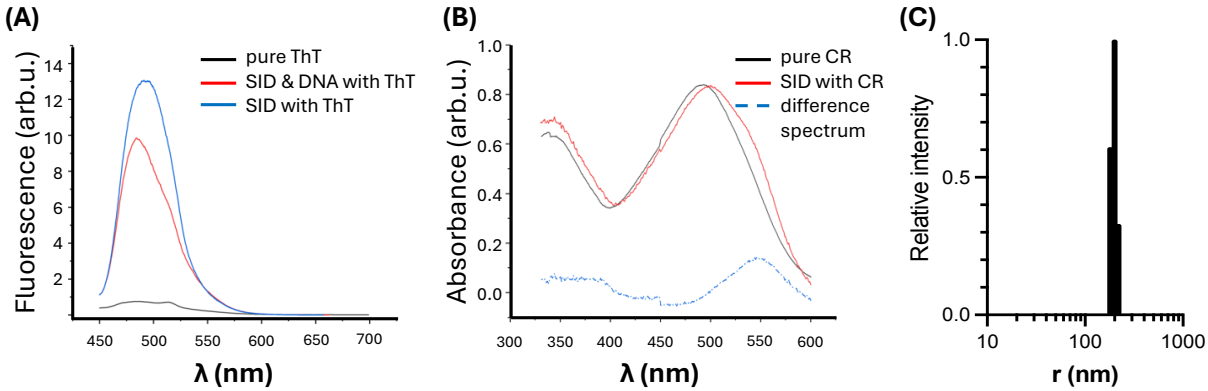

Figure 2. (A) - fluorescence spectrum of Thioflavin T when adding the model SID peptide to a solution; (B) absorption spectra of Congo red and its complexes with the SID peptide (C) Histogram of aggregate size distribution in fibril solutions. The distribution of scattering intensity is given

According to atomic force microscopy, the average fibril length was 500±50 nm, structural data obtained from cryo-electron microscopy indicate that the fibril consists of 12 protofibrils (Fig. 1B). The distance between the beta sheets is 2.5Å, that is, a protofibril with a length of around 500 nm represents about 2000 molecules stacked in parallel. If a fibril consists of a bundle of 12 protofibrils, then it contains 24,000 monomer molecules (1338.81 Da). Together, these data allow us to estimate the molecular weight of one fibril as around 32000 kDa.

The calculated [44] hydrodynamic radius of fibrils (in the elongated particle model) was about 150 nm, thus the distribution of radii in the solution (Fig.2C) corresponds to the model obtained using transmissive microscopy.

3.2 Peptide SID is capable of forming complexes with short model DNA

The model DNAs used were 45 bp dsDNA (1P) labeled at the 3' end with Cy5 (Table 1). The sequence of the model dsDNA 1P was chosen based on (i) the absence of self-complementarity, (ii) the high (80°C) melting temperature of double-stranded DNA, and (iii) the high content of GC pairs.

The presence of the binding was determined by DNA retardation in agarose gel. Using serial dilutions of peptide fibrils, the minimum concentration of the peptide at which no DNA entry into the gel was observed was determined (Figure 3). For 10 nM 1P DNA, this peptide

concentration was 300 μM. Based on the results of the experiment, the ratio of the number of peptide molecules to the number of DNA molecules was calculated, which was 1 DNA molecule per 30,000 peptide molecules, which, according to structural data, approximately corresponds to 1 DNA molecule per 1 fibril.

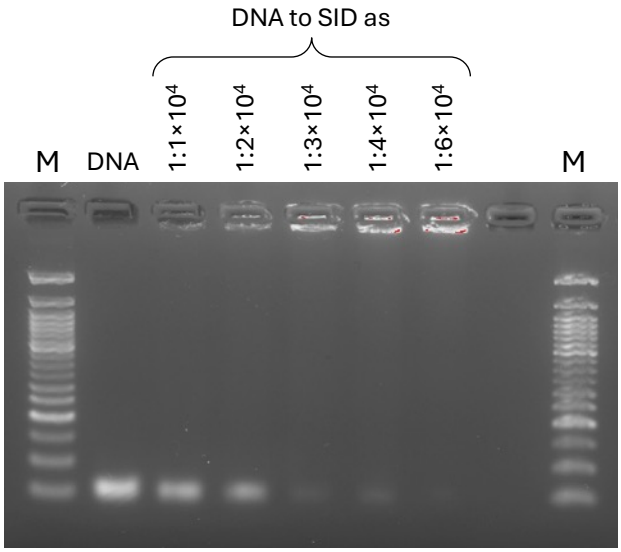

Figure 3. Titration of DNA 1P solution with SID peptide. M – 1000 kbase marker, «DNA» - DNA 1P 10nM, DNA:SID ratios used are designated above lanes

The study of the morphology of DNA/peptide complexes by atomic force and electron microscopy, as well as by dynamic light scattering, did not show detectable differences between in morphology and hydrodynamic radii in such samples. Considering that the length of 45 bp of DNA is about 17 nm (versus 500 nm of fibril length), it is possible to assume the presence of structural features of the fibril, for example, arising at its ends and allowing DNA binding. If the binding of short DNA occurred at the strand of the fibril, the structural identity of the repeating elements would lead to a greater number of bound DNA molecules per fibril. It should be also noted that when using unlabeled DNA, 1P fibrils showed deterioration in binding to Thioflavin T (Figure 2A, red line), which may be a consequence of the coincidence of binding sites for DNA and dye on the fibril and competition between them.

3.3 DNA-peptide equilibrium dissociation constant estimation

Microscale thermophoresis experiments were performed to determine the equilibrium dissociation constant and assess the specificity of peptide binding to model DNA. Sixteen samples were prepared, each containing a constant concentration of fluorescent DNA and concentrations of peptide ranging from 0.021 nM to 700 nM.

The data presented (Figure 4) show that when the peptide concentration reaches 5.4 nM, peptide-DNA complexes are formed. The calculated equilibrium dissociation constant KD was about 10^{-7} M, which corresponds to highly specific binding.

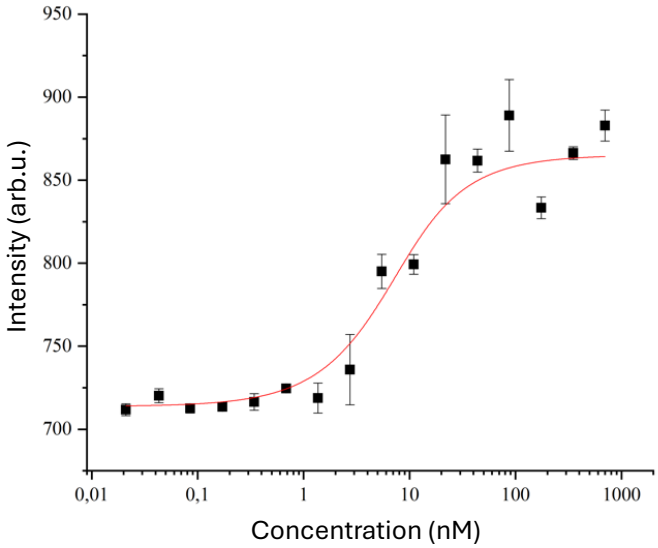

Figure 4. Plot of intensity versus concentration of SID peptide in the sample.

3.4 Mechanisms of interaction between the peptide and model DNA

To determine possible mechanisms of binding of the peptide to the model DNA, experiments on molecular dynamics simulation in the free diffusion mode were carried out. Figures 5 and 6 show the results of the analysis of the data obtained as a result of the simulation. As an example, Fig. 6 shows a heat map of the distances between nucleotides of the model DNA for two of the 90 peptides used in the simulation.

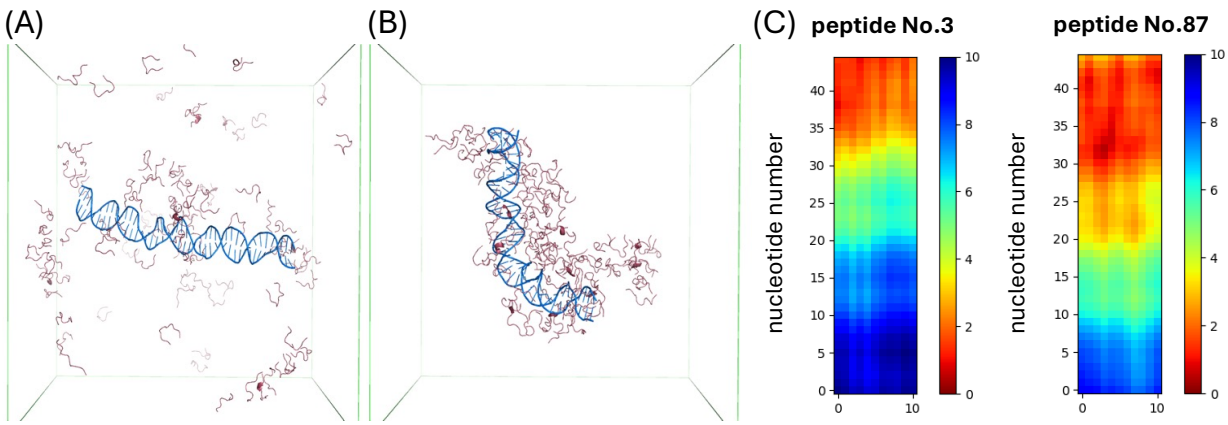

Figure 5. Sequential changes occurring to the system at 0 ns (A) and 150 ns (B) stages.

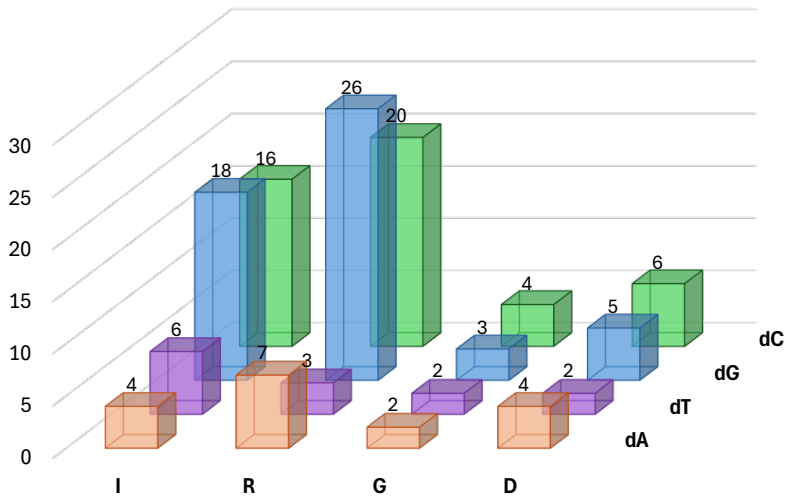

Figure 6. Maps of the modes of distances of amino acid residues of peptides SID No. 3(a) and No. 87(b) to the nitrogenous bases of model DNA 1P, during the simulation.

Table 2. Number of contacts of nucleotides with amino acid residues of the peptide during the simulation process

	I	R	G	D
dA	4	7	2	4
dT	6	3	2	2
dG	18	26	3	5
dC	16	20	4	6

Based on the data presented in Table 2, a hypothesis was put forward about the selective interaction of the arginine and isoleucine residues of the SID peptide with GC-rich regions of DNA 1P.

3.5 Interaction of fibrils with phage lambda DNA

To test the hypothesis about the interaction of peptide molecules in the fibrils with GC-pairs of DNA, the lambda phage DNA was chosen as a model object. Conditions were selected for the formation of complexes of the SID peptide with the lambda phage DNA for subsequent restriction analysis using restriction enzymes that recognize GC- and AT-rich regions.

In experiments on retardation with lambda phage DNA (1 nM), the maximum concentration of the peptide that completely prevents DNA from entering the gel was 0.6 μM

(Figure 7, 1:6*10^5 ratio). This ratio was subsequently used to carry out the enzymatic hydrolysis reaction of DNA on the DNA-peptide complex.

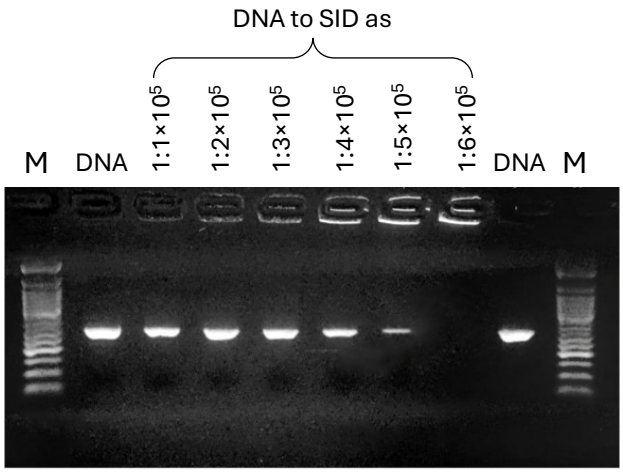

Figure 7. - Titration of lambda phage DNA solution with SID peptide. «M» -molecular weight marker 1000base, «DNA» -lambda phage DNA 1 nM, DNA:SID ratios used are designated above lanes

It should be noted that preliminary experiments to study the effect of concentrations of SID peptide comparable to concentrations of restriction endonucleases showed that it is not capable of directly inhibiting the enzymatic activity of endonucleases (data not shown).

To determine the binding sites of SID peptide to phage λ DNA, three different enzymes were selected: Ava1(CYCGRG), HindIII(AAGCTT) and Nde1(CATATG), catalyzing the reaction of DNA hydrolysis by specific sites noted in brackets.

The results of electrophoretic analysis of the products of the reaction of DNA hydrolysis of phage lambda are presented in Figure 8

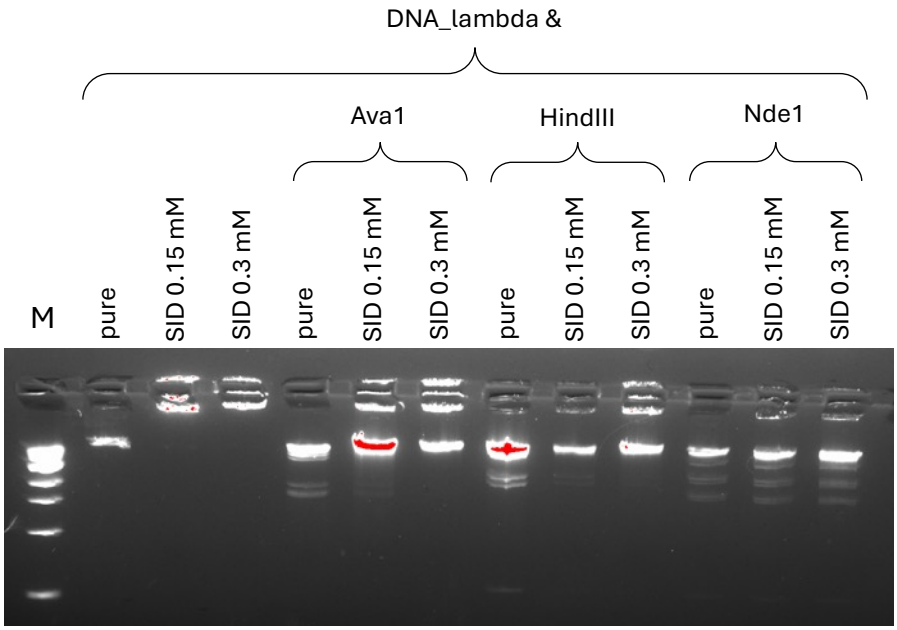

Figure 8. Result of electrophoretic separation of lambda phage DNA segments obtained using restriction enzymes Ava1, HindIII, Nde1 in the presence of SID. «M» – 1000kb marker, "pure" – phage λ DNA enzymes used and SID concentrations used are designated above lanes

The obtained results allow us to assume that inhibition of the hydrolysis reaction occurs only for enzymes that have specificity for recognizing GC-rich (Ava1(CYCGRG), HindIII(AAGCTT)) but not AT-rich regions (Nde1(CATATG)), which is consistent with the data obtained as a result of computer modeling and indicates the presence of selectivity in the binding of the peptide to the dsDNA molecule, depending on the nucleotide composition. Apparently, the peptide bound to DNA prevents the formation of the enzyme-substrate complex and protects the phosphodiester backbone of DNA from hydrolysis catalyzed by the enzyme.

Peculiarities of SID Peptide Binding to Short and Long DNA

For lambda phage DNA, the ratio preventing DNA from entering the gel is about $6*10^5$ peptide molecules per 1 DNA molecule, or about 20 fibrils per 1 DNA molecule, while for 1P DNA this ratio was 1 fibril per 1 DNA molecule. Comparison of the length of lambda phage DNA (about 16 μm) and SID fibril (about 0.5 μm) suggests that lateral interaction of fibrils and DNA may be involved in binding of lambda phage DNA, in contrast to binding to short DNA, where fibril ends are apparently take part in. We hypothesized that different binding mechanisms may be involved in the interaction of DNA of different lengths with peptide, which are, however, based on the same mechanism of specific interaction with GC-rich regions. To test this hypothesis, an experiment was conducted to titrate already formed complexes of the SID peptide/λ phage DNA

with short 1P DNA followed by retardation in a gel. To conduct the experiment, the minimum concentration of the SID peptide was selected, leading to a complete delay of the λ phage DNA at the entrance to the gel, the concentrations of 1P DNA corresponded to those used earlier, during retardation of SID/1P complexes (Figure 9).

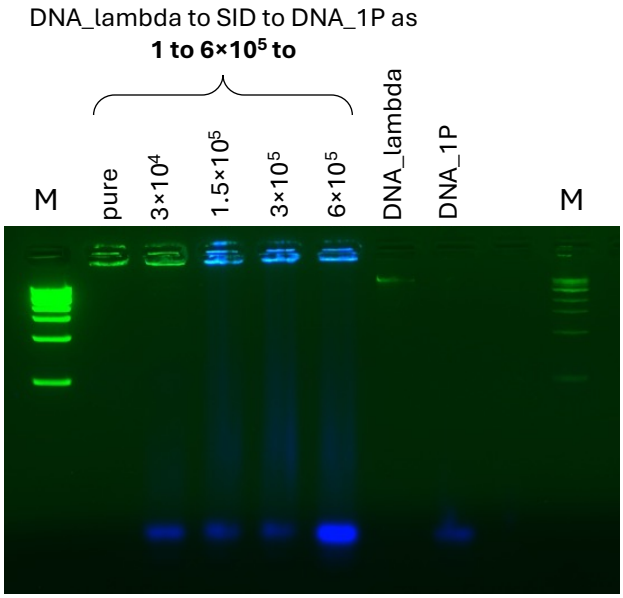

Figure 9. - Titration of SID – phage λ DNA complexes with 1P DNA solution. «M» - 1000base marker, «pure» - complex of lambda phage DNA with SID peptide $1/6*10^5$ SID concentrations used are designated above lanes

The results of the experiment showed that the capacity of peptide complexes for 1P DNA does not depend on whether it is in a complex with lambda phage DNA, i.e., the binding of long and short DNA by the SID peptide possibly occurs independently.

To compare structures of the DNA-fibrils complexes and peptide fibrils alone on small scales, small-angle neutron scattering (SANS) experiments were carried out on complexes of SID peptide fibrils with lambda phage DNA and 1P DNA. The obtained SANS spectra are presented in Figure 10.

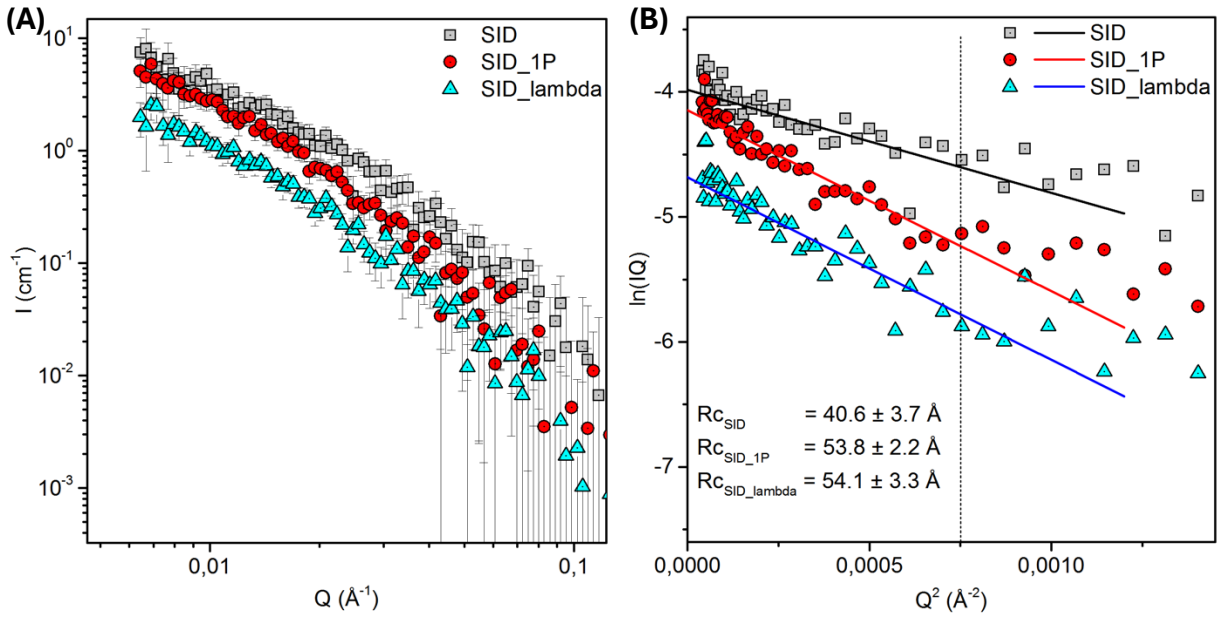

Figure 10. (A) Small-angle neutron scattering spectra in logarithmic coordinates. I is the scattering intensity, q is the scattering vector (B) SANS spectra on solutions of SID peptide, SID + 1P DNA, SID + phage lambda DNA, plotted in lnIq coordinates from q^2 to determine the thickness of fibrillar structures. The dots show the experimental data, the lines show the linear approximation by the least squares method in the Guinier region (q*Rc < 1.3).

Based on the nature of the spectra obtained, as well as on previous results obtained by the AFM method, it was concluded that in solution, the SID peptide itself and in combination with various types of dsDNA forms filamentous complexes of different radii.

Thus, the average radius of cylindrical structures formed in solutions of SID peptide and DNA of different lengths and nucleotide compositions was obtained (Figure 10 (B), Table 3). The obtained data satisfied the Guinier condition: q·R_C<1.3.

Table 3. Calculated values of the cylindrical complexes radii

Object	Calculated radius of cylinder
SID	Rc = 40.6 ± 3.7
SID + 1P	Rc = 53.8 ± 2.2
SID + λ	Rc = 54.1 ± 3.3

The average radius of the cylindrical complexes formed as a result of the interaction of the SID peptide with DNA 1P and phage λ differed from the average radius of the cylindrical complexes formed by the peptide itself by 10 Å, which is explained by the presence of DNA

molecules in the complexes, which leads to an increase in the diameter of the cylindrical aggregates formed by the peptides by 20 Å.

The results of the experiment also showed that the structures of the complexes of lambda phage DNA with a peptide and 1P DNA with a peptide do not differ on scales of 6–60 nm, corresponding to the detectable range of SANS. It should be taken into account that the length of 1P DNA (17 nm) is less than the persistent length of DNA (45–70 nm, depending on the ionic strength), i.e. the difference in the binding mechanisms of 50 bp DNA and lambda phage DNA manifests itself on scales exceeding the persistent length and is associated with additional degrees of freedom arising due to the mobility of long DNA.

Conclusions

We have shown that the IRIDIRIDIRG peptide carrying an ionic self-complementary motif is capable of forming amyloid-like fibrils. The peptide fibrils interact with a model 45 bp DNA, the equilibrium dissociation constant of such a complex is about 10^{-7} M. Although the mechanisms of interaction of the peptide fibrils with a short model DNA (45 bp) and a long model DNA (about 50 thousand nucleotide pairs) differ, the structural basis for the interaction of the peptide fibrils with DNA are GC pairs in DNA with Arginine residues in the peptide. The resulting complexes can be used both for DNA delivery and as a basis for peptide drugs that are potentially capable of regulating gene expression.

Acknowledgements

The results of the work were obtained using computational resources of the supercomputer center in Peter the Great Saint-Petersburg Polytechnic University Supercomputing Center (www.spbstu.ru), SANS facilities of Joint Institute for Nuclear Research (project no. 2020-10-20-21-14-42), with financial support of National Research Center "Kurchatov Institute" (order No. 2557, October 28, 2021) and is partially funded by the Ministry of Science and Higher Education of the Russian Federation as part of World-class Research Center program: Advanced Digital Technologies (contract No. 075-15-2022-311 dated by 20.04.2022). MST experiments were supported by the Russian Science Foundation grant No. 22-14-00278 (to A.L.K.).

References

[1] A. Heguy, A. A. Stewart, J. D. Haley, D. E. Smith, and J. G. Foulkes, "Gene Expression as a Target for New Drug Discovery," 1995.

[2] S. A. Lambert *et al.*, "The Human Transcription Factors," Feb. 08, 2018, *Cell Press*. doi: 10.1016/j.cell.2018.01.029.

[3] N. M. Luscombe, S. E. Austin, H. M. Berman, and J. M. Thornton, "An overview of the structures of protein-DNA complexes," *Genome Biol*, vol. 1, no. 1, pp. 1–37, 2000.

[4] L. Marcos-Villar, A. Pazo, and A. Nieto, "Influenza Virus and Chromatin: Role of the CHD1 Chromatin Remodeler in the Virus Life Cycle," *J Virol*, vol. 90, no. 7, pp. 3694–3707, 2016, doi: 10.1128/jvi.00053-16.

[5] A. V. Vasin, O. A. Temkina, V. V. Egorov, S. A. Klotchenko, M. A. Plotnikova, and O. I. Kiselev, "Molecular mechanisms enhancing the proteome of influenza A viruses: An overview of recently discovered proteins," Jun. 24, 2014, *Elsevier*. doi: 10.1016/j.virusres.2014.03.015.

[6] J. Pei *et al.*, "Nuclear-localized human respiratory syncytial virus NS1 protein modulates host gene transcription," *Cell Rep*, vol. 37, no. 2, p. 109803, 2021, doi: 10.1016/j.celrep.2021.109803.

[7] T. Morii and K. Makino, "Sequence-specific DNA binding by short peptides," *Advances in DNA Sequence-Specific Agents*, vol. 4, no. C, pp. 105–137, 2002, doi: 10.1016/S1067-568X(02)80007-X.

[8] V. Khavinson *et al.*, "Neuroprotective effects of tripeptides—epigenetic regulators in mouse model of Alzheimer's disease," *Pharmaceuticals*, vol. 14, no. 6, 2021, doi: 10.3390/ph14060515.

[9] F. Chiti and C. M. Dobson, "Protein misfolding, functional amyloid, and human disease," *Annu Rev Biochem*, vol. 75, pp. 333–366, 2006, doi: 10.1146/annurev.biochem.75.101304.123901.

[10] M. S. Rubel *et al.*, "Functional mammalian amyloids and amyloid-like proteins," Sep. 01, 2020, *MDPI AG*. doi: 10.3390/life10090156.

[11] D. Otzen and R. Riek, "Functional amyloids," *Cold Spring Harb Perspect Biol*, vol. 11, no. 12, Dec. 2019, doi: 10.1101/cshperspect.a033860.

[12] E. M. Y. Moresco, D. La Vine, and B. Beutler, "Prion-like behavior of MAVS in RIG-I signaling.," *Cell Res*, vol. 21, no. 12, pp. 1643–1645, Dec. 2011, doi: 10.1038/cr.2011.155.

[13] L. A. Ivanova *et al.*, "Matrix is everywhere: extracellular DNA is a link between biofilm and mineralization in Bacillus cereus planktonic lifestyle," *NPJ Biofilms Microbiomes*, vol. 9, no. 1, Dec. 2023, doi: 10.1038/s41522-023-00377-5.

[14] Z. Du, K. W. Park, H. Yu, Q. Fan, and L. Li, "Newly identified prion linked to the chromatin-remodeling factor Swi1 in Saccharomyces cerevisiae," *Nat Genet*, vol. 40, no. 4, pp. 460–465, Apr. 2008, doi: 10.1038/ng.112.

[15] Y. Wang, N. Zolotarev, C. Y. Yang, A. Rambold, G. Mittler, and R. Grosschedl, "A Prion-like Domain in Transcription Factor EBF1 Promotes Phase Separation and Enables B Cell Programming of Progenitor Chromatin," *Immunity*, vol. 53, no. 6, pp. 1151-1167.e6, Dec. 2020, doi: 10.1016/j.immuni.2020.10.009.

[16] M. Yamada, S. Hara, T. Yamada, F. Katagiri, K. Hozumi, and M. Nomizu, "Double-stranded DNA stereoselectively promotes aggregation of amyloid-like fibrils and generates peptide/DNA matrices," in *Biopolymers - Peptide Science Section*, John Wiley and Sons Inc, Nov. 2014, pp. 465–472. doi: 10.1002/bip.22571.

[17] G. D'Auria *et al.*, "Self-assembling properties of ionic-complementary peptides," *J Pept Sci*, vol. 15, no. 3, pp. 210–219, Mar. 2009, doi: 10.1002/psc.1083.

[18] M. Kabiri, I. Bushnak, M. T. McDermot, and L. D. Unsworth, "Toward a mechanistic understanding of ionic self-complementary peptide self-assembly: role of water molecules and ions," *Biomacromolecules*, vol. 14, no. 11, pp. 3943–3950, Dec. 2013, doi: 10.1021/bm401077b.

[19] M. Altman, P. Lee, A. Rich, and S. Zhang, "Conformational behavior of ionic self-complementary peptides," *Protein Sci*, vol. 9, no. 6, pp. 1095–1105, Jun. 2000, doi: 10.1110/ps.9.6.1095.

[20] V. V. V Egorov, Ya. A. Zabrodskaya, D. V. V Lebedev, A. N. N. Gorshkov, and A. I. I. Kuklin, *Structural features of the ionic self-complementary amyloidogenic peptide*, vol. 848, no. 1. 2017, p. 12022. doi: 10.1088/1742-6596/848/1/012022.

[21] Z. Wan, “Arginine-Rich Ionic Complementary Peptides and Their Drug Delivery Potential,” 2013.
[22] “Preparation of samples of DNA molecules for AFM measurements.” [Online]. Available: www.ntmdt.ru
[23] D. Nečas and P. Klapetek, “Gwyddion: An open-source software for SPM data analysis,” *Central European Journal of Physics*, vol. 10, no. 1, pp. 181–188, 2012, doi: 10.2478/s11534-011-0096-2.
[24] H. LeVine, “Thioflavine T interaction with synthetic Alzheimer’s disease beta-amyloid peptides: detection of amyloid aggregation in solution.,” *Protein Sci*, vol. 2, no. 3, Mar. 1993, doi: 10.1002/pro.5560020312.
[25] Y. A. Y. A. Zabrodskaya *et al.*, *The amyloidogenicity of the influenza virus PB1-derived peptide sheds light on its antiviral activity*, vol. 234, no. January. [Online]. Available: https://doi.org/10.1016/j.bpc.2018.01.001
[26] A. G. Soloviev, T. M. Solovjeva, O. I. Ivankov, D. V. Soloviov, A. V. Rogachev, and A. I. Kuklin, “SAS program for two-detector system: Seamless curve from both detectors,” *J Phys Conf Ser*, vol. 848, no. 1, 2017, doi: 10.1088/1742-6596/848/1/012020.
[27] Schrödinger LLC, “The PyMOL Molecular Graphics System, Version 2.4.”
[28] M. J. Abraham *et al.*, “GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers,” *SoftwareX*, vol. 1–2, pp. 19–25, Sep. 2015, doi: 10.1016/J.SOFTX.2015.06.001.
[29] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling, “ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB,” *J Chem Theory Comput*, vol. 11, no. 8, pp. 3696–3713, Jul. 2015, doi: 10.1021/acs.jctc.5b00255.
[30] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, “Comparison of simple potential functions for simulating liquid water,” *J Chem Phys*, vol. 79, no. 2, pp. 926–935, Jul. 1983, doi: 10.1063/1.445869.
[31] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak, “Molecular dynamics with coupling to an external bath,” *J Chem Phys*, vol. 81, no. 8, pp. 3684–3690, 1984, doi: 10.1063/1.448118.
[32] S. Nosé, “A unified formulation of the constant temperature molecular dynamics methods,” *J Chem Phys*, vol. 81, no. 1, pp. 511–519, Jul. 1984, doi: 10.1063/1.447334.
[33] W. G. Hoover, “Canonical dynamics: Equilibrium phase-space distributions,” *Phys Rev A (Coll Park)*, vol. 31, no. 3, pp. 1695–1697, Mar. 1985, doi: 10.1103/PhysRevA.31.1695.
[34] G. J. Martyna, M. L. Klein, and M. Tuckerman, “Nosé–Hoover chains: The canonical ensemble via continuous dynamics,” *J Chem Phys*, vol. 97, no. 4, pp. 2635–2643, Aug. 1992, doi: 10.1063/1.463940.
[35] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, “Explicit reversible integrators for extended systems dynamics,” *Mol Phys*, vol. 87, no. 5, pp. 1117–1157, Apr. 1996, doi: 10.1080/00268979600100761.
[36] M. Parrinello and A. Rahman, “Polymorphic transitions in single crystals: A new molecular dynamics method,” *J Appl Phys*, vol. 52, no. 12, pp. 7182–7190, Dec. 1981, doi: 10.1063/1.328693.
[37] S. Nosé and M. L. Klein, “Constant pressure molecular dynamics for molecular systems,” *Mol Phys*, vol. 50, no. 5, pp. 1055–1076, Dec. 1983, doi: 10.1080/00268978300102851.
[38] Schrödinger LLC, “The PyMOL Molecular Graphics System, Version~1.8,” Nov. 2015.
[39] D. Tegunov and P. Cramer, “Real-time cryo-electron microscopy data preprocessing with Warp,” *Nat Methods*, vol. 16, no. 11, pp. 1146–1152, 2019, doi: 10.1038/s41592-019-0580-y.
[40] T. Wagner *et al.*, “SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM,” *Commun Biol*, vol. 2, no. 1, p. 218, 2019, doi: 10.1038/s42003-019-0437-z.

[41] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination," *Nat Methods*, vol. 14, no. 3, pp. 290–296, 2017, doi: 10.1038/nmeth.4169.

[42] Y. A. Zabrodskaya, A. V. Shvetsov, V. B. Tsvetkov, and V. V. Egorov, "A double-edged sword: supramolecular complexes of triazavirine display multicenter binding effects which influence aggregate formation," Aug. 13, 2019, *Taylor and Francis Ltd.* doi: 10.1080/07391102.2018.1507837.

[43] G. Lancelot, R. Mayer, and C. Hélène, "Models of interaction between nucleic acids and proteins. Hydrogen bonding of arginine with nucleic acid bases, phosphate groups and carboxylic acids," *Biochim Biophys Acta*, vol. 564, no. 2, pp. 181–190, Sep. 1979, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/486477

[44] "HR Calculator," https://calculator.academy/protein-hydrodynamic-radius-calculator/.