Journal of Molecular Liquids xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Novel insights on nucleopeptide binding: A spectroscopic and *in silico* investigation on the interaction of a thymine-bearing tetrapeptide with a homoadenine DNA

Domenica Musumeci [a,b], Samee Ullah [c], Aamer Ikram [c], Giovanni N. Roviello [b,*]

[a] Department of Chemical Sciences, University of Naples Federico II, Via Cintia 21, 80126 Naples, Italy
[b] Institute of Biostructures and Bioimaging (IBB) – CNR, 80145 Naples, Italy
[c] Department of Virology, National Institute of Health, Park Road, Chak Shahzad, 45500 Islamabad, Pakistan

ARTICLE INFO

Article history:
Received 6 August 2021
Revised 12 October 2021
Accepted 25 October 2021
Available online xxxx

Keywords:
Nucleopeptide
DNA binding
CD spectroscopy
Molecular modelling
CD simulation
computational CD interpretation
DichroCalc

ABSTRACT

Nucleopeptides are a class of molecules with numerous applications in the field of therapy, diagnostics and biomaterials development. Despite their nucleobase-decorated nature, their binding to natural nucleic acid targets does not necessarily involve all nucleopeptide bases, as we showed in this study. Here, we present a CD study on the interaction of a dithymine-functionalized tetra-L-serine with a homoadenine DNA (dA_{12}) reporting an interpretation of the experimental data in light of our computational studies based on molecular docking and molecular dynamics (MD), as well as computer-assisted CD interpretation and simulation of the predicted complex structure. The stoichiometry of the complex, emerged by CD titration, accounted for a 1:2 T:A ratio. Hence, we supposed that binding did not involve a full pairing of the complementary bases but a partial thymines engagement. This hypothesis was sustained by the docking and MD simulations performed on the selected ligand and the complementary target of DNA and RNA, used for comparison. The nucleopeptide bound the DNA through a single A-T recognition involving complementary base-pairing, as well as by some interactions between its backbone (and in particular L-serine OH) and the nucleic acid. Overall, this confirmed that nucleopeptides can interact with nucleic acids leaving some of their nucleobases free for establishing further interactions with other biomolecules or for crosslinking in supramolecular structures in aqueous solution. Nevertheless, even though no typical DNA secondary structure is formed after nucleopeptide-binding, this ligand is able to induce a higher degree of structuration in the random deoxyoligonucleotide target as evidenced by CD, MD and CD simulation.

1. Introduction

Nucleopeptides [1–9] are a class of chimeric compounds bearing DNA nucleobases connected to peptide backbones. They can play an important role in biomedical research not only for their potential as nucleic acid-targeting molecules [10–14] but also because, similarly to peptides, [15–20] they show protein-binding ability and can modulate protein-associated biological pathways [21,22]. Moreover, nucleopeptides can form supramolecular structures, held together by both base-base pairing and other non-covalent molecular recognitions [23–26]: possibly the nucleopeptide-assembly can be reinforced by the interaction with synthetic ligands [27], aspect of clear utility in nanobiotechnological applications. In this regard, even molecules containing single nucleobases, such as synthetic nucleoside analogues [28–34] nucleoamino acids [35,36] and peptidyl nucleosides [37–40] are interesting structures, some of which were shown to act as building blocks for supramolecular networks of potential biomedical relevance [41–47].

Despite their promising properties, nucleopeptide structures are still in need of deeper investigation; indeed, even if a certain number of examples of nucleobase-bearing peptides able to hybridize DNA and RNA targets are known [48–51], the exact role of nucleobases in complementary nucleic acid binding is an aspect not yet fully clarified. In this context, computational studies can facilitate the comprehension of the interaction mechanism at the basis of DNA or RNA-nucleopeptide recognition [52,53].

We have previously found that a nucleopeptide whose backbone was based on L-serine (Fig. 1), endowed with particularly interesting self-assembling properties, was able to interact with a

* Corresponding author.
E-mail addresses: giroviel@unina.it, giovanni.roviello@cnr.it (G.N. Roviello).

https://doi.org/10.1016/j.molliq.2021.117975
0167-7322/

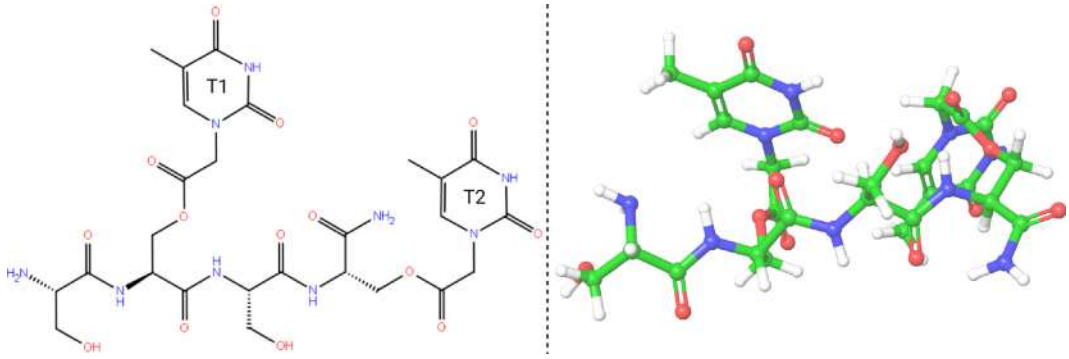

Fig. 1. Structural representation (left) of the dithymine tetrapeptide (with the N-terminal thymine labelled as *1*) and 3D view (right) of the molecule after energy minimization achieved by LigPrep and visualization in Maestro (Schrodinger).

complementary DNA affording, anyway, only minor changes in the structure of the nucleic acid target, as revealed by CD spectroscopic experiments [23]. The L-serine tetrapeptide backbone carried two unfunctionalized amino acids, introduced as spacers between the two nucleobase-bearing moieties in order to mimic the base-base distance in natural DNA, following the scheme of nucleopeptide design proposed by Diederichsen [54]. Thus, we expected that, besides the base-complementarity, the overall interaction with DNA and RNA could be reinforced by H-bonding involving the free hydroxyl groups on the nucleopeptide and chemical functionalities on the nucleic acid target.

Herein, we investigated in more detail by CD experiments the behaviour of the dithymine L-serine tetrapeptide in order to achieve a quantitative information on its complexes with DNA. Moreover, we performed several computational studies in order to give an interpretation to the experimental results and better describe the binding modality.

2. Materials and methods

The complete workflow of the methodology of the present study is provided in Fig. S1.

2.1. Chemicals

The Fmoc-L-Ser(tBu)-OH and TCH_2COOH were obtained from Sigma Aldrich (Milan, Italy). DNA A_{12} were purchased from Biomers (Ulm, Germany).

2.2. Nucleopeptide synthesis

The dithymine tetraserine nucleopeptide (Fig. 1) was synthesized in solid phase on a Rink-amide resin using the commercially available Fmoc-L-Ser(tBu)-OH and the *ad hoc* synthesized thyminyl-L-serine monomer according to a procedure described previously by some of us [23]. After deprotection and detachment from the solid support, the nucleopeptide was purified by reverse-phase HPLC and its identity was confirmed by comparison with the literature data [23].

2.3. Spectroscopic studies

Circular dichroism (CD) spectra were obtained with procedures similar to previous literature reports [55–59] on a Jasco J-810 spectropolarimeter, equipped with a Peltier ETC-505 T temperature controller, using a Hellma-238-QS tandem quartz cell (2 × 0.437 5 cm, Fig. 2) and a 1-cm cell (Fig. 3). UV absorption spectra have been collected simultaneously to CD on the same instrument [60] to minimize errors induced by separate measurements on different instruments [61].

2.4. In silico studies

2.4.1. Structure preparation

The A_{12} single strand DNA and RNA oligomers were constructed in Maestro, Schrödinger 2021-1 (Schrödinger Release 2021-1: Maestro, Schrödinger, LLC, New York, NY, 2021). The preparation wizard was used to add the hydrogens atoms, and the structure was optimized using default parameters *via* the OPLS3e force field [62]. The side chains atoms were added, and the bond orders, overlapping atoms, alternate position of the atoms were corrected through the build panel. Lastly, the steric clashes and the improper torsions were checked before the docking experiment [63].

2.4.2. Receptor grid generation and XP-docking

The entire nucleic acid (DNA and RNA) receptor carrying a total of 12 nucleotides was selected to generate the grid *via* the receptor grid generation panel using default parameters. The grid was refined by the ligand diameter midpoint box set to 15 Å for all three axes. The nucleopeptide ligand was minimized through LigPrep ligand preparation panel using the default parameters (Schrödinger Release 2021-1: LigPrep, Schrödinger, LLC, New York, NY, 2021). The energetically-minimized structure of the nucleopeptide was subjected to the Glide-XP (Extra-Precision algorithm) docking engine [64]. The ligand sampling was set to flexible.

2.4.3. Molecular dynamics simulation protocol

The dA_{12} and the nucleopeptide-nucleic acid (DNA and RNA) complexes were subjected to Desmond simulation engine for 100 ns [65]. The Intel® Core™ i9-9900 K Desktop CPU (central processing unit) with 64-GB of RAM on Linux Fedora 34 Scientific (64bit) OS was used for the simulations and the analysis.

2.4.4. System building

The default predefined TIP3P solvent model was constructed for the nucleopeptide-nucleic acid (DNA and RNA) complexes using the OPLS3e force field [66]. The box boundary conditions were set to cubic with distance of x y z coordinates set to 10 Å. The calculated volume of the box was 151324 and 145015 $Å^3$ for DNA and RNA complexes, respectively. The salt concentration was set to 0.1 M. The neutralization of the system was achieved by adding 11 and 12 Na ions for DNA and RNA, respectively. The full system contained 14160 and 13392 atoms for DNA and RNA complexes, respectively.

2.4.5. Trajectory production and analysis

The simulation time was set to 100 ns and the ensemble class of NPT was selected. Temperature was set to 300 K and pressure to 1 atm. The default protocol "relax model before simulation algorithm" was selected and the trajectories were produced for the 100 ns. Finally, the 100 ns trajectories were analyzed for the RSMD (Root Mean Square Deviation).

2.4.6. Binding free energy calculations

The MMGBSA algorithm of the prime module was used to better explore the interactions of the nucleopeptide in the pocket of DNA and RNA receptor. The complexes nucleopeptide-nucleic acid (DNA and RNA) were subjected to the MM-GBSA module and the ligand binding free energy was thus, calculated.

2.4.7. DNA secondary structure prediction

CD spectral data related to the experiment described in Fig. 2 were extracted in the 230–315 nm wavelength range and used to create the input files that were uploaded in CD-NuSS [67] with which the secondary structure predictions were run for both DNA and nucleopeptide-DNA complex.

2.4.8. CD predictions

Structure PDB files for dA_{12} and its complex with the nucleopeptide obtained after the MD simulations were processed by DichroCalc [68], a web interface for protein circular and linear dichroism calculations that we used for CD DNA simulations after manual PDB file editing replacing the unrecognized 'DA' text for adenine-containing deaoxyribonucleotides with 'A'. The PDB files for both dA_{12} and nucleopeptide-DNA complex were then uploaded as input files in DichroCalc obtaining finally the predicted CD files which were edited with SpectraGryph 1.2 [69].

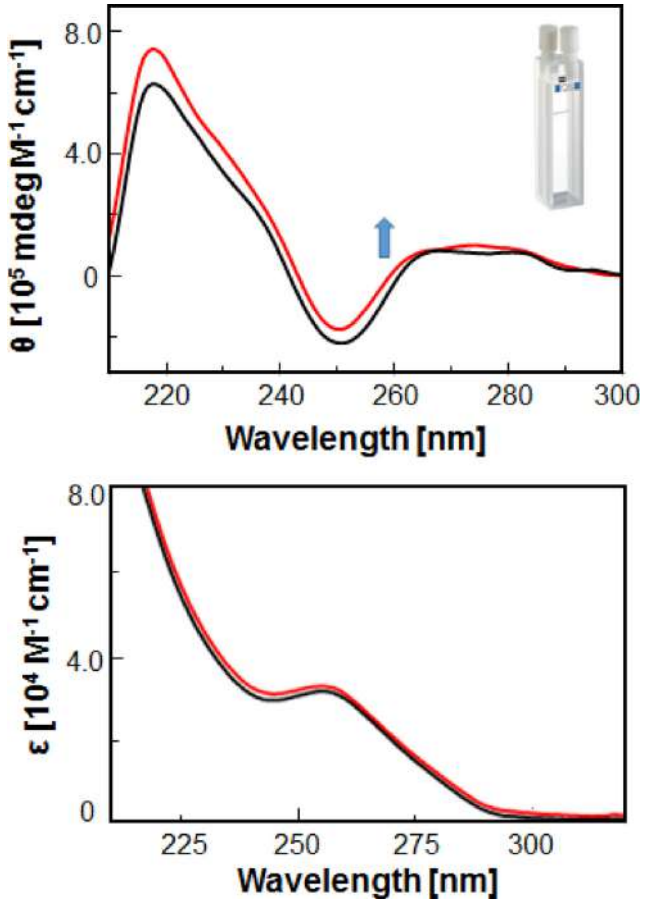

Fig. 2. CD (up) and UV (down) spectra before (black) and after (red) mixing of nucleopeptide with dA_{12} (T:A = 1:1, 15 μM) in 10 mM phosphate buffer (pH 7.5) at 5 °C in a tandem dual quartz cell. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Results and discussion

3.1. CD and UV binding studies

To study the nucleopeptide interaction with the single-stranded DNA, we used as target a dA_{12} oligonucleotide, and recorded simultaneously CD and UV spectra of both DNA and ligand solutions, placed in the two separated chambers of a tandem dual quartz cell, before and after manual rotation of the cell (Fig. 2). This particular methodology permits to run experiments minimizing concentration errors and is used in binding experiments with chiral ligands [70,71].

Only a minor change in the aromatic region of the CD spectrum could be observed after complexation accompanied by a slight blue shift (Fig. 2, up). Interestingly, no evident hypochromism was detected in UV absorbance spectra after mixing nucleopeptide and DNA solutions (Fig. 2, bottom) suggesting that no significant base-base stacking occurred as one would have expected in the case of W-C base pairing.

The same experiment previously performed by us with a RNA target of the same sequence of the DNA had not resulted in any CD change, which had led us to conclude that the nucleopeptide was unable to establish significant interactions with RNA [23].

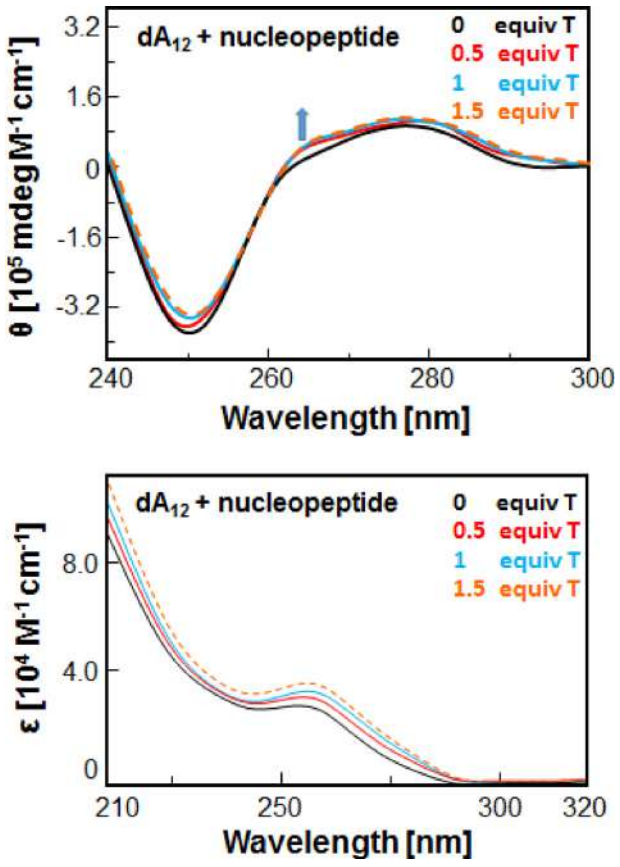

Fig. 3. CD (up panel) and UV (down) titrations of dA_{12} in the presence of 0–1.5 equivalents in base of nucleopeptide ligand at 5 °C in 10 mM phosphate buffer (pH 7.5, 1 cm quartz cell).

3.2. CD- and UV-monitored titration

To further explore the DNA-binding ability evidenced by the above-described experiment, and add a quantitative information on the stoichiometry of the DNA-nucleopeptide complex, we performed the CD/UV-monitored titration reported in Fig. 3.

Remarkably, by titrating the DNA solution with nucleopeptide, only few changes in the CD spectral curve of the nucleic acid were detected in the aromatic spectral region (Fig. 3, up), taking into account that at the used concentrations the nucleopeptide does not contribute significantly to the CD signal. A slight blue shift can be recognized for the DNA-nucleopeptide complex and a substantial stabilization of the signal occurred in CD spectra after adding nucleopeptide at a T:A = 1:2 ratio (Fig. 3, up). This seemed to indicate that the nucleopeptide molecule interacted with the adenine-rich DNA strand without pairing the two thymine bases but only involving a single T:A pairing. During titration, the nucleopeptide addition provoked a progressive increase in UV absor-

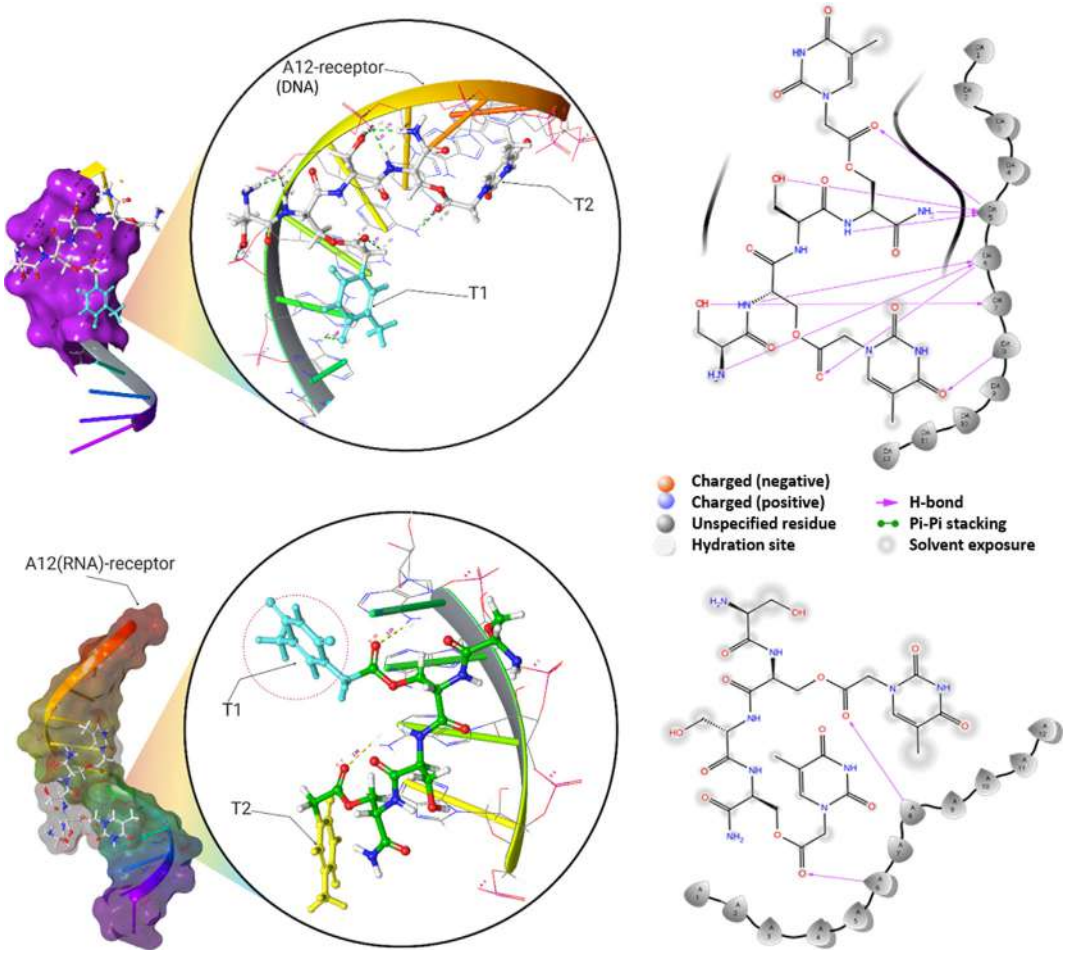

Fig. 4. Upper: 3D interactions of the nucleopeptide (T1 ring in cyan color) with dA_{12} receptor after molecular docking (left panel); the green dotted lines represent the H-bonds. 2D interactions diagram of the ligand with the dA_{12} receptor residues (right); the pink arrow lines represent the H-bonds. Bottom: 3D interaction of the ligand with the A_{12} (RNA) receptor after docking (left panel). The T1 is colored in cyan and T2 in yellow. The yellow dotted lines represent the H-bonds interactions. 2D interaction diagram of A_{12} (RNA)-nucleopeptide complex (right). The pink arrow lines represent the H-bonds. The A6 and A8 are in light grey color. (For interpretation of the references to color in this figure legend, the reader should consider the web version of this article.)

Table 1
XP H-bond interaction energy analysis of nucleopeptide-dA_{12} and nucleopeptide A_{12} RNA (last two lines) complexes with interactions (donor or acceptor), distance (Å) and energies (kcal/mol) indicated.

Receptor-atoms	Ligand-atoms	Interaction	Distance (Å)	Energy (kcal/mol)
dA5-H:468	O:132	H-donor	1.72	−6.7
dA5-H:449	O:132	H-donor	1.68	−3.4
dA5-H:445	O:132	H-donor	1.73	−6.3
dA5-O:393	H:152	H-acceptor	1.91	−3.2
dA6-O:164	H:461	H-donor	1.73	−6.6
dA6-O:164	H:464	H-acceptor	1.75	−0.7
dA6-O:410	H:184	H-donor	2.31	−2.2
dA7-H:467	O:196	H-donor	1.75	−3.2
dA8-H:247	O:417 (T1 ring)	H-acceptor	2.03	−0.9
RA6-H:190	O:7	H-donor	2.08	−2.7
RA8-H:256	O:24	H-donor	2.23	−1.1

bance at about 260 nm perfectly matching the expected thymine UV-contribution (Fig. 3, bottom). This confirmed the absence of any significant stacking event involving nucleopeptide thymines and DNA adenine bases during the binding — which instead would have caused a lower increase than the observed one — as already revealed by the previous assay (Fig. 2, bottom).

3.3. DNA-nucleopeptide molecular modelling

To achieve a more accurate interpretation of the experimental findings, and possibly of the binding modality for the interaction of the nucleopeptide with the dA_{12} target, we performed computational studies using molecular docking and dynamics, as well as *in silico* methods for CD interpretation and simulation. In particular, initial docking experiments were achieved using the DNA as receptor and the nucleopeptide as ligand.

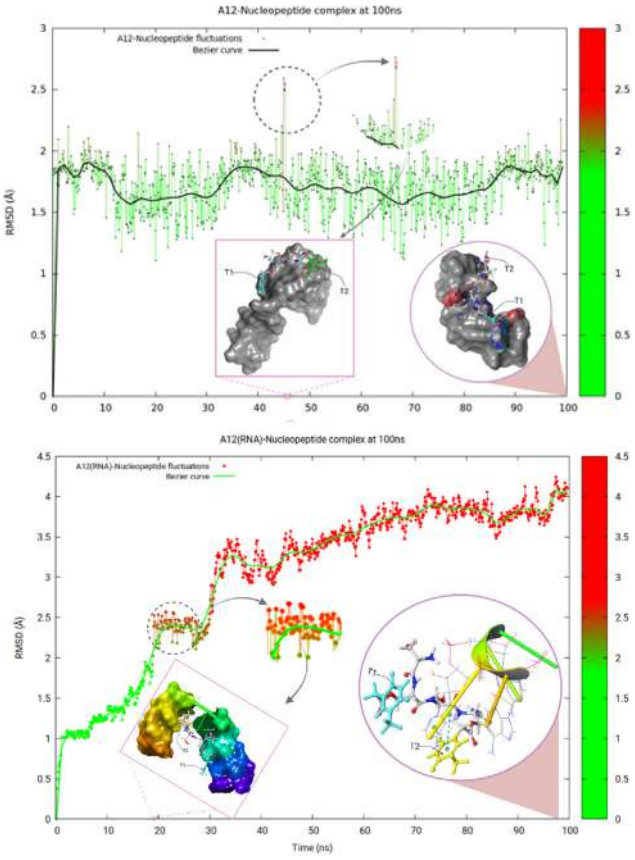

Fig. 5. Upper: root mean square deviation (RMSD) of the nucleopeptide-dA_{12} complex as a function of time (ns). The right-hand side scale shows the RMSD values in the 0–2 Å (with green color corresponding to the most stable system) and 2–3 Å (with red color indicating the least stable system) ranges. The interaction of the T2-ring (green colored) can be observed at 45 ns corresponding to a fluctuation of RMSD reaching 2.6 Å. The interaction of the T1-ring (cyan colored) at 100 ns trajectory point with RMSD less than 2 Å can also be observed, validating the potential binding mode of the nucleopeptide *via* T1 ring and not T2 to the dA_{12} receptor through the 100 ns simulations. Bottom: RMSD of the nucleopeptide-A12 RNA complex as a function of time (ns). (For interpretation of the references to color in this figure legend, the reader should consider the web version of this article.)

3.3.1. Nucleopeptide-A_{12} DNA and RNA docking

We investigated the interactions of the nucleopeptide in proximity of the dA_{12} *via* Glide XP-docking program (upper Fig. 4, left panel). In our simulation, the nucleopeptide bound the dA_{12} target with a predicted Glide XP (Extra-Precision) docking energy of −34 kcal/mol. The nucleopeptide interacted with the dA_{12} receptor *via* a network of nine hydrogen bonds and more in detail through a four H-bonds interaction with A5, three H-bonds with A6 and a single H-bond with A7 at distances of 1.72, 1.68, 1.73, 1.91, 1.73, 1.75, 2.31 and 1.75 Å respectively. Furthermore, the T1 base interacted by H-bonding with A8 at a 2.03 Å distance (Fig. 4, right). We performed a similar docking study also on A_{12} RNA (as a control target) using the same software and conditions used for DNA (Fig. 4, bottom). The nucleopeptide interacted with the RNA with a predicted Glide XP-docking (Extra-Precision) energy of −4.1 kcal/mol. Interestingly, the ligand interacted with the A_{12} RNA receptor via a network of two hydrogen bonds involving the A6-O2-atom, and A8-O2-atom, with 2.08 Å and 2.23 Å distances, respectively (Fig. 4, bottom). Interestingly, compared to DNA the molecular docking analysis of the nucleopeptide in the pocket of the RNA receptor revealed lower-level interactions, in terms of H-bonding networks and XP energies, which explains that the ligand binds more efficiently to DNA.

3.3.2. XP H-bond interactions energy analysis

To achieve a better comprehension of the binding mode of the nucleopeptide toward the DNA, we performed XP H-bond interactions energy analysis by glide algorithm module. The XP H-bond interaction energy module of XP-docking technique identifies the crucial residues of the receptor involved in binding with the ligand and calculates their hydrogen bonding energies in kcal/mol. In our case, the highest H-bond energies were obtained for the A5, A6 and A7 atoms of the receptor, whilst for A8 we found the lowest energy (−0.9 kcal/mol, Table 1).

Similarly, we performed XP H-bond interactions energy analysis via glide algorithm module also for RNA binding. The H-bond energies recorded for the RA6 and RA8 atoms of the receptor (RNA) are shown in the Table 1. Comparing XP H-bond energies of RNA and DNA complexes, we observed higher-level interactions, in terms of H-bonding networks and energies, in the DNA complex with respect to the RNA one. To further validate the docking and XP H-bond interaction energy analysis, we explored both the nucleic acid complexes in a 100 ns simulation as described in the protocol present in the methodology section.

3.3.3. Conformational dynamics of nucleopeptide-dA_{12} complex

The ligand/nucleic acid (DNA and RNA) complex conformational stability was assessed by 100 ns of molecular dynamics simulations and the resulting trajectories were analyzed for the structure parameter i.e. the root mean square deviation (RMSD). Fig. 5

Table 2
Nucleopeptide/DNA H-bond interactions, distance (Å) and energies (kcal/mol) after 100 ns molecular dynamics simulation.

Receptor-atoms	Ligand-atoms	Interactions	Distance (Å)	Energy (kcal/mol)
A6-O:164	H:467	H-acceptor	1.75	−0.8
A6-H:461	O:164	H-donor	2.31	−3.7
A6-H:183	O:420	H-donor	2.50	−3.4
A12-H:375	O:417	H-acceptor	2.30	−0.6
A12- N:367	H:460	H-donor	2.09	−4.1

depicts the RMSD changes against the trajectory for the DNA complex showing that the complex attained the structural stability prior to (remaining stable until) the end of the MD simulation. The full system was well equilibrated and a RMSD lower than 2.5 Å over the entire 100 ns trajectory for the complex indicates a stable system during our simulation.

A fluctuation of RMSD reaching a value of 2.6 Å at 45 ns (evidenced in Fig. 5) corresponded to the interactions of the T2 with the nucleic acid. However, from 45 ns onwards (45–100 ns) the interactions of the T2 were replaced by T1 with the DNA target.

Interestingly, the interaction of T1 already evidenced in docking (Fig. 4) which was based on the nucleopeptide interaction with A8 *via* a single H-bond of energy −0.9 (kcal/mol) was confirmed and explored in details by molecular dynamics simulations revealing more potent interactions with two H-bonds of −4.1 and −0.6 kcal/-mol (with A12) and a single one (with A6) of −0.8 kcal/mol energy with the DNA (Table 2). The nucleopeptide also interacted with A6 through further two H-bonds of −3.7 (kcal/mol) and −3.4 (kcal/-mol) energy but without involving any thymine rings. Furthermore, comparing the interactions of the docking complex to molecular dynamics simulations we observed the loss of the A5, A7 and A8 interactions with the nucleopeptide probably due to the movement of the T1-ring towards the A12 and the lower H-bond energies especially of A8 with T1. However, the interactions of the A6 were retained except the loss of a H-bond from a NH_2 group. Overall, these *in silico* findings suggest that the binding mode of the nucleopeptide involves a single thymine (and more likely T1 ring than T2) in the recognition of the DNA target receptor, a hypothesis validated by us *via* the 100 ns molecular dynamic simulations as well as our experimental data (Figs. 3, 5 and 6).

Owing to the simulation on RNA complex, Fig. 5 (bottom) depicts the RMSD changes against the trajectory showing that the complex attained the structural stability from 0 to 18 ns (recorded RMSD 2.5 Å). However, a trend of fluctuations from 19 ns (RMSD 2.6 Å) highlighted in red–orange color to onwards 98 ns (RMSD 4.3 Å) at the end of the MD simulation shed light on the structural instability of the RNA complex. Extracting the frames at 19 (Fig. S2) and 98 (Fig. 6) ns revealed that the ligand interacted with RA7 and RA5 via double π-π stacking interactions involving the ring centers (RCEN) of RA7-T2 (highlighted in yellow) with distances RCEN:1-RCEN:2=4.31 Å, RCEN:1-RCEN:3=4.13 Å, and RA5-T2 with distances RCEN:1-RCEN:2=4.10 Å, RCEN:1-RCEN:3=2.9 Å (Figs. S2, 6). Interestingly, no similar π-π stacking interactions were revealed in the DNA throughout the 100 ns simulations.

3.3.4. Simulation quality analysis

The simulation quality module of Maestro program was utilized for a simulation quality analysis and descriptive parameters such as the total energy, potential energy, temperature, pressure and volume of the full system were extracted as described for the DNA complex in Fig. 7 and Table 3. The simulation quality analysis revealed an average total energy of −34790.990 (kcal/mol), potential energy −41956.933 (kcal/mol), temperature 298.748(K), pressure 1.452(bar) and volume of 119911.166 $Å^3$. In the case of the RNA complex, the simulation quality analysis showed an average total energy of −25050.764 (kcal/mol), potential energy −36083.491 (kcal/mol), temperature 297.672 (K), pressure 0.180 (bar) and volume of 166627.571 $Å^3$.

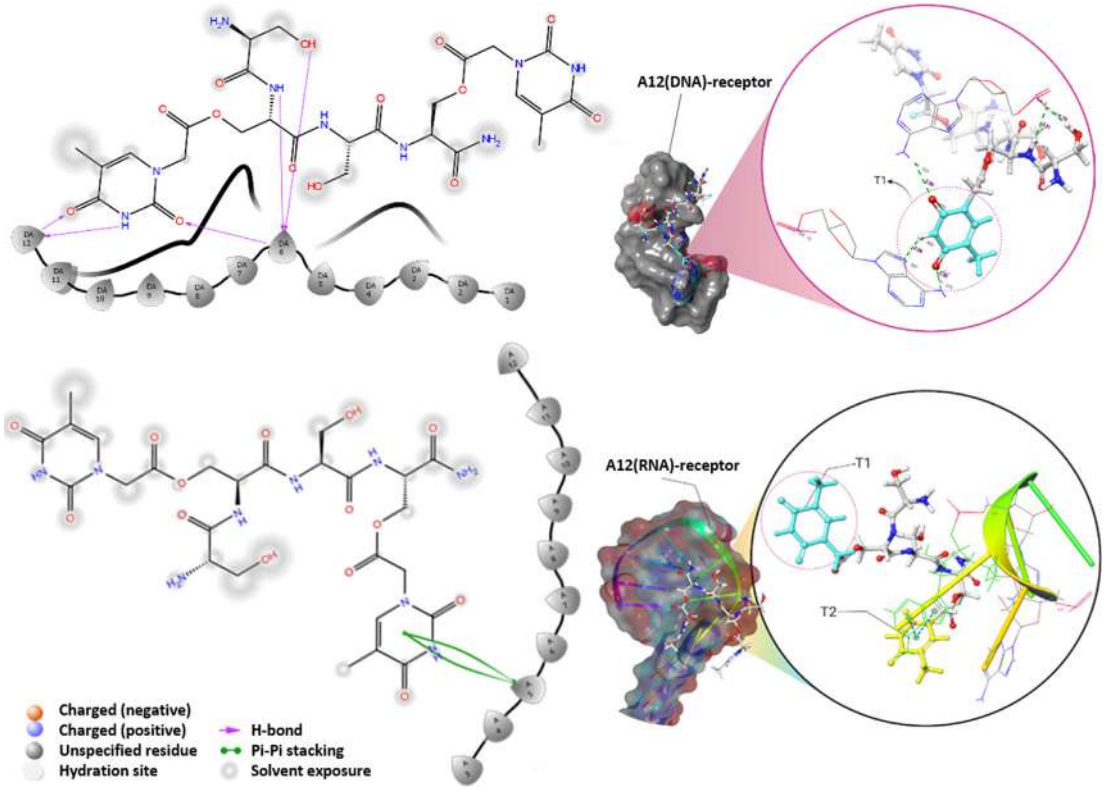

Fig. 6. Upper: nucleopeptide-dA_{12} complex 2D interaction (pink arrows represents the H-bonds) at 100 ns trajectory (left panel). 3D interactions of the nucleopeptide (T1 ring highlighted in circle and colored in cyan) at 100 ns (right). Bottom: 2D (left) and 3D (right) interactions of the nucleopeptide-RNA complex at 98 ns. The green lines represent the π-π stacking interactions (left). The T2 is highlighted in yellow and T1 in cyan color, while the RA5 in green color (right). (For interpretation of the references to color in this figure legend, the reader should consider the web version of this article.)

Interestingly, only one complementary nucleobases interaction was revealed by MD confirming what already suggested by our spectroscopic studies described in Fig. 3, indicating a 1:1 = T:A binding stoichiometry in the case of homoadenine DNA (Fig. 3) for the serine-based nucleopeptide, that, on the other hand, was unable to interact significantly with a RNA of same sequence [23].

3.3.5. Ligand binding free energy calculations

The ligand binding free energy was estimated for both DNA and RNA complexes *via* prime MMGBSA (Molecular Mechanics/Generalized Born Model and Solvent Accessibility) using the OPLS3e force field, VSGB2.1 solvent model (Schrödinger Release 2021-1: Maestro, Schrödinger, LLC, New York, NY, 2021). The Glide XP docking complexes from the pose-viewer were subjected to the prime MM-GBSA simulation module and the total binding free energy was predicted. The MMGBSA predicts the approximate binding free energy and the more negative energy indicates clearly the better binding. The formula to calculate the ΔG_{Bind} *via* the prime MMGBSA (Molecular Mechanics/Generalized Born Model and Solvent Accessibility) is the following.

$$\Delta G_{Bind} = G_{Complex} - G_{Receptor} - G_{Ligand}$$

The MMGBSA ΔG_{Bind} values were computed as −28.62 and −3.10 kcal/mol for the DNA and RNA complexes, respectively, indicating a higher binding affinity of the dithymine tetrapeptide for the DNA target. The more negative ΔG_{Bind} value found for the DNA complex compared to the RNA one, furnishes more insights on the nucleopeptide binding to the DNA receptor which was also investigated in a 100 ns molecular dynamics simulation. The higher level molecular interactions of the nucleopeptide, in terms of H-bonding networks, XP-energies, less fluctuations and higher stability in the conformational dynamics compared to RNA, all concur to define more favourable interaction characteristics of the dithymine tetrapeptide in the vicinity of the DNA when compared to ribonucleic targets in agreement with our experimental findings [23].

3.3.6. CD data interpretation by in silico methods

Possible secondary structures of the DNA-nucleopeptide complex were investigated using CD-NuSS [67], a web-based program

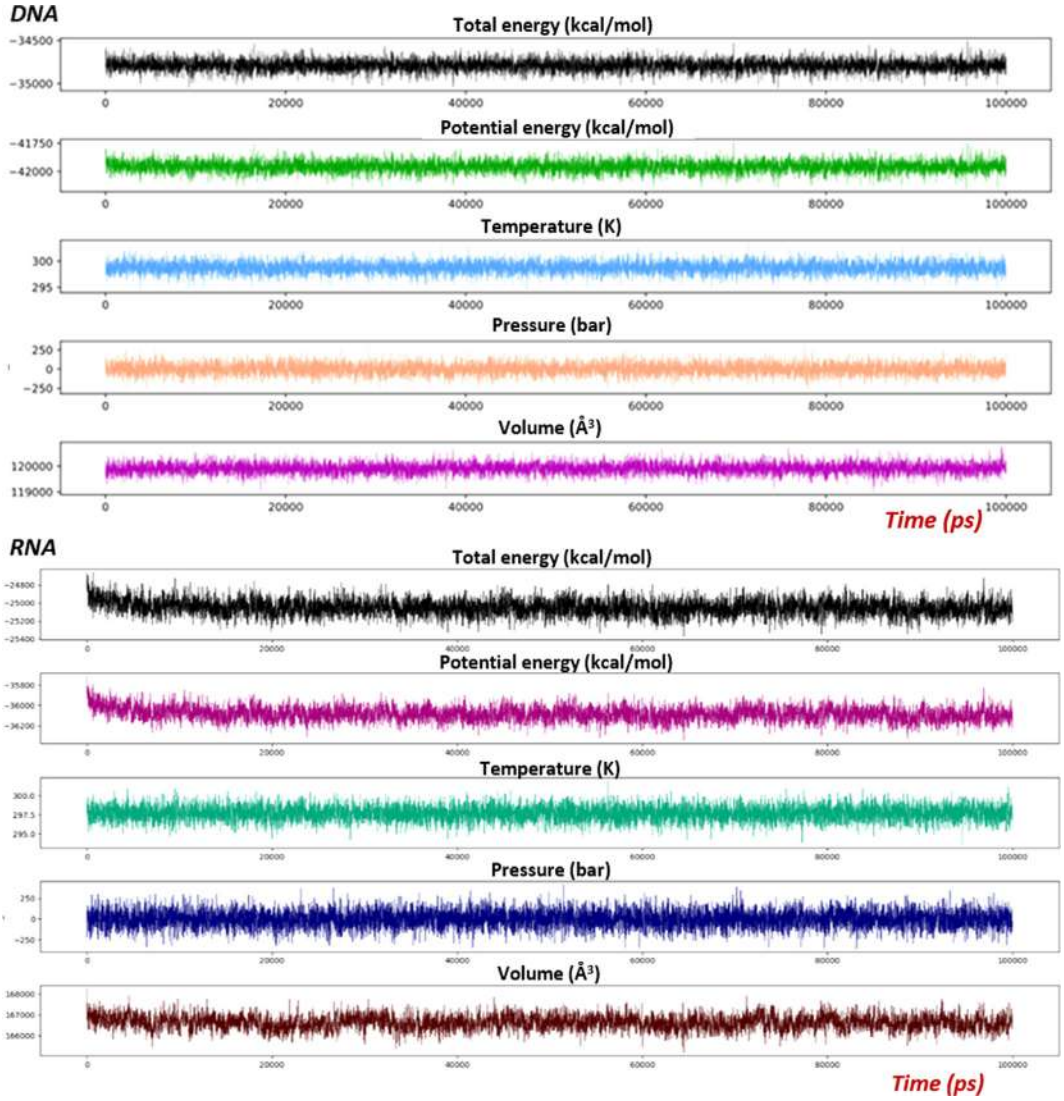

Fig. 7. Upper: descriptive parameters for the nucleopeptide-dA_{12} system at 100 ns. E = Total energy (kcal/mol) in black color, EP = Potential energy (kcal/mol) in green, T = Temperature (K) in blue, P = Pressure (bar) in light orange and V = Volume (Å^3) in violet. Bottom: Descriptive parameters of the nucleopeptide-A_{12} (RNA) system at 100 ns. E = Total energy (kcal/mol) in black color, EP = Potential energy (kcal/mol) in purple, T = Temperature (K) in light green, P = Pressure (bar) in navy and V = Volume (Å^3) in brown color. (For interpretation of the references to color in this figure legend, the reader should consider the web version of this article.)

Table 3
Simulation quality analysis of the full system at 100 ns for the nucleopeptide complexes with DNA and RNA. Slope (ps^{-1}) was 0 in all cases.

	DNA		RNA	
Extracted MD properties	Average	Std. Dev.	Average	Std. Dev.
Total energy (kcal/mol)	−34790.990	68.765	−25050.764	88.096
Potential energy (kcal/mol)	−41956.933	52.198	−36083.491	69.124
Temperature (K)	298.748	1.167	297.672	0.979
Pressure (bar)	1.452	82.365	0.180	102.581
Volume ($Å^3$)	119911.166	214.194	166627.571	341.259

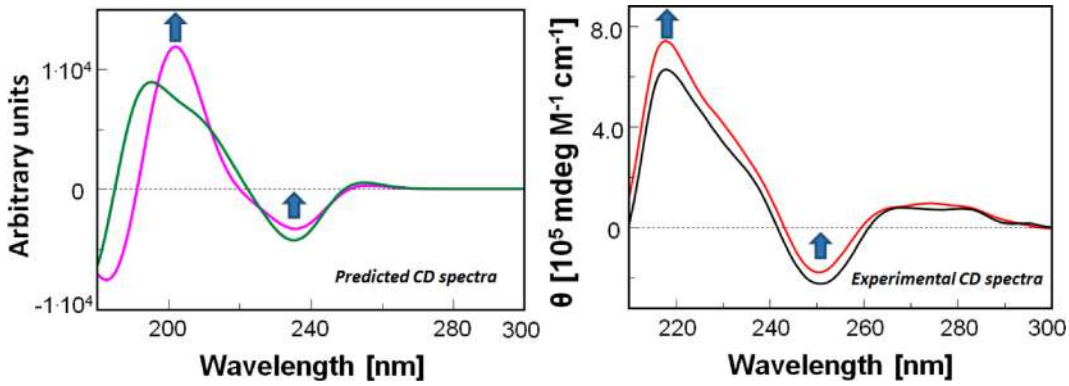

Fig. 8. Theoretical CD spectra for dA_{12} (green) and its complex with the nucleopeptide (magenta) as simulated by DichroCalc [68] starting from the PDB structure files predicted by our computational modelling studies (left panel). Experimental CD spectra before (black) and after (red) mixing of the nucleopeptide with dA_{12} (T:A = 1:1, 15 μM) in 10 mM phosphate buffer (pH 7.5) at 5 °C in a tandem dual quartz cell (right panel). (For interpretation of the references to color in this figure legend, the reader should consider the web version of this article.)

which makes use of eXtreme Gradient Boosting (XGBoost) algorithm to predict the nucleic acid secondary structures which correspond to the CD spectral datasets furnished as inputs. Among the several nucleic acid structures that one can identify by CD-Nuss, there are DNA A-, B- and Z-form, triplex (parallel and antiparallel), G-quadruplex (parallel, antiparallel, and hybrid), RNA stem-loop, DNA-RNA duplex, etc. [67]. Thus, we elaborated the specific data corresponding to the CD spectral images of Fig. 2 and the corresponding numerical X-(CD) and Y-(wavelength) axes values, extracted in the range between 230 and 315 nm, were furnished to CD-NuSS. As expected, no DNA secondary structure was predicted in the case of the unbound single stranded dA_{12}, clearly in the randomly coiled state. Owing to its complex with the nucleopeptide, CD-NuSS indicated that the interaction of the thymine-bearing tetra-L-serine with the DNA did not induce any typical DNA secondary structures (Fig. S3).

Furthermore, the predicted model of the DNA-nucleopeptide complex emerged from the above-described molecular dynamics was used to simulate its CD spectrum by DichroCalc [68]. In particular, a significant increase in the CD intensity in the <230 nm wavelength region was predicted by DichroCalc (Fig. 8, left panel), which is to some extent in analogy to what we experimentally found by experimental CD (Fig. 8, right) indicating a higher structuration degree in the random-coil DNA caused by the interaction with the nucleopeptide. An intensity decrease in CD spectrum minima at 240–250 nm were observed in both predicted and experimentally recorded spectra (Fig. 8). As reported in the literature [72,73] the <230 nm band arises from secondary structural DNA features, whereas the aromatic band (>260 nm) probes the 3D conformation of DNA such as bending and curvature. Thus, our experimental and predictive CD studies seem to indicate that the nucleopeptide binds DNA leading to appreciable secondary structure variations, that can be associated with the single AT pairing revealed by CD titration and confirmed by MD simulation, but without significant effects on bending and curvature of the nucleic target.

4. Conclusions

Herein we investigated the nucleic acid binding behaviour of a dithymine tetra-L-serine nucleopeptide by means of spectroscopic and *in silico* methods. We found that the nucleopeptide was able to form complexes with DNA based on a 2:1 = A:T stoichiometric ratio, as experimentally revealed by our CD titration experiment. Complexation assays confirmed the ability of the nucleopeptide to cause slight changes in the DNA secondary structure corresponding mainly to an increase of the <230 nm positive band and, to a lesser extent, to a blue shift in the CD band of the complex, while no significant hypochromism was detected in the UV absorbance as effect of the nucleopeptide/DNA molecular recognition, suggesting that adenine bases were not significantly stacked with thymine counterparts. *In silico* studies conducted by docking the nucleopeptide to the DNA target object of the spectroscopic experiments, and subsequent MD simulations, suggested that only T1 (H-bonding with A12 in an almost perpendicular AT interaction similarly to other aromatic rings found in biological structures [74]), as well backbone (SER1 NH and OH) moieties of the nucleopeptide are likely involved in DNA binding leaving the second base (T2) potentially free to form, for example, supramolecular assemblies able to incorporate hydrophobic drugs as previously demonstrated for this molecule [23] and other nucleopeptides [75]. RNA docking and dynamic studies were also conducted in order to obtain a comparison between the predicted nucleopeptide binding modes occurring in the case of DNA with respect to RNA. Overall, these simulations revealed with the ribonucleic target a different interaction mode based on π-π stacking that led to a complex less stable than that predicted by us for DNA, confirming our

previously published results [23] on the higher affinity of the serine-based nucleopeptide for deoxyoligonucleotide targets than RNAs. Taken together, all these findings led us to hypothesize that the dodecadenine DNA is bound by multiple nucleopeptide units through single not W-C AT base pairs as well H-bonding involving the nucleopeptide backbone and the DNA bases and phosphate moieties. This occurred despite the nucleobase-containing peptide having been designed according to the general rule of a DNA mimicking nucleopeptide whose repeating unit is built by two α-amino acid moieties, one of which derivatized with the nucleobase, directly joined to one another by a peptide bond [54], which revealed in RNA binding in several previous literature reports [12,48,49,76]. Overall, we presented for the first time a combined spectroscopic and computational/CD-predictive approach which enabled us to i) give an interpretation of the structure of the complex formed in our spectroscopic experiments, confirming the lack of any typical DNA secondary structure element; ii) validate indirectly the computed structure model, emerged by MD for the complex, comparing the predicted binding-associated CD changes (including CD intensity increase for the positive <230 nm band) with those experimentally found, iii) conclude that despite being unable to induce any typical DNA secondary structure in its complementary nucleic target, the nucleopeptide provoked in the random dA_{12} a conformational rearrangement involving a higher structuration degree determined by the interaction of its backbone and one of its nucleobases, with the nucleic acid.

5. Authors' contribution

GNR designed research, wrote part of the paper, performed experiments, computed theoretical CD and gave their interpretations; DM performed experimental research, analyzed the data and wrote part of the paper; SU performed docking and MD studies and wrote part of the paper; AI contributed to study preparation and docking and MD data interpretation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Dr. Antonietta Gargiulo (IBB-CNR) for her technical assistance, Dr. Elliot Drew (Queen Mary University of London), Dr David M. Rogers (University of Nottingham), Patrik Nikolić and Dr. Vedran Miletic (RxTx Research, www.rxtx.tech/research, Rijeka, Croatia) for useful discussion on computational chemistry methods.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molliq.2021.117975.

References

[1] G.N. Roviello, M. Moccia, R. Sapio, M. Valente, E.M. Bucci, M. Castiglione, C. Pedone, G. Perretta, E. Benedetti, D. Musumeci, J. Pept. Sci. 12 (2006) 829.
[2] G.N. Roviello, E. Benedetti, C. Pedone, E.M. Bucci, Amino Acids 39 (2010) 45.
[3] G.N. Roviello, D. Musumeci, M. Moccia, M. Castiglione, R. Sapio, M. Valente, E. M. Bucci, G. Perretta, C. Pedone, Nucleosides Nucleotides Nucleic Acids 26 (2007) 1307.
[4] G.N. Roviello, D. Musumeci, C. Pedone, E.M. Bucci, Amino Acids 38 (2008) 103.
[5] P. Geotti-Bianchini, A. Moretto, C. Peggion, J. Beyrath, A. Bianco, F. Formaggio, Org. Biomol. Chem. 8 (2010) 1315.
[6] P. Geotti-Bianchini, J. Beyrath, O. Chaloin, F. Formaggio, A. Bianco, Org. Biomol. Chem. 6 (2008) 3661.
[7] P. Geotti-Bianchini, M. Crisma, C. Peggion, A. Bianco, F. Formaggio 611 (2009) 37.
[8] P. Geotti-Bianchini, M. Crisma, C. Peggion, A. Bianco, F. Formaggio, Chem. Commun. (2009) 3178.
[9] R.A. Kramer, K.H. Bleicher, H. Wennemers, Helv. Chim. Acta 95 (2012) 2621.
[10] G. Roviello, D. Musumeci, M. Castiglione, E.M. Bucci, C. Pedone, E. Benedetti, J. Pept. Sci. 15 (2009) 155.
[11] G.N. Roviello, D. Musumeci, V. Roviello, M. Pirtskhalava, A. Egoyan, M. Mirtskhulava, Beilstein J. Nanotechnol. 6 (2015) 1338.
[12] G.N. Roviello, D. Musumeci, RSC Adv. 6 (2016) 63578.
[13] T. Takahashi, K. Hamasaki, A. Ueno, H. Mihara, Bioorg. Med. Chem. 9 (2001) 991.
[14] H. Miyanishi, T. Takahashi, H. Mihara, Bioconjug. Chem. 15 (2004) 694.
[15] F. Viparelli, A. Cassese, N. Doti, F. Paturzo, D. Marasco, N.A. Dathan, S.M. Monti, G. Basile, P. Ungaro, M. Sabatella, C. Miele, R. Teperino, E. Consiglio, C. Pedone, F. Beguinot, P. Formisano, M. Ruvo, J. Biol. Chem. 283 (2008) 21769.
[16] S. La Manna, C. Di Natale, D. Florio, D. Marasco, Int. J. Mol. Sci. 19 (2018) 2714.
[17] S. La Manna, L. Lopez-Sanz, S. Bernal, S. Fortuna, F.A. Mercurio, M. Leone, C. Gomez-Guerrero, D. Marasco, Eur. J. Med. Chem. 221 (2021) 113547.
[18] P. Scognamiglio, C. Di Natale, G. Perretta, D. Marasco, Curr. Med. Chem. 20 (2013) 3803.
[19] C. Di Natale, S. La Manna, I. De Benedictis, P. Brandi, D. Marasco, Front. Pharmacol. 11 (2020).
[20] A. Falanga, P. Melone, R. Cagliani, N. Borbone, S. D'Errico, G. Piccialli, P. Netti, D. Guarnieri, Molecules 23 (2018) 1655.
[21] G.N. Roviello, S. Di Gaetano, D. Capasso, A. Cesarani, E.M. Bucci, C. Pedone, Amino Acids 38 (2009) 1489.
[22] G.N. Roviello, S. Di Gaetano, D. Capasso, S. Franco, C. Crescenzo, E.M. Bucci, C. Pedone, J. Med. Chem. 54 (2011) 2095.
[23] G.N. Roviello, D. Musumeci, E.M. Bucci, C. Pedone, Mol. BioSyst. 7 (2011) 1073.
[24] X. Li, Y. Kuang, H.C. Lin, Y. Gao, J. Shi, B. Xu, Angew. Chem. Int. Ed. 50 (2011) 9365.
[25] D. Yuan, X. Du, J. Shi, N. Zhou, J. Zhou, B. Xu, Angew. Chem. 127 (2015) 5797.
[26] A.D. Noblett, K. Baek, L.J. Suggs, ACS Biomater. Sci. Eng. 7 (2021) 2605.
[27] G.N. Roviello, G. Roviello, D. Musumeci, E.M. Bucci, C. Pedone, Amino Acids 43 (2012) 1615.
[28] G. Oliviero, J. Amato, N. Borbone, S. D'Errico, G. Piccialli, E. Bucci, V. Piccialli, L. Mayol, Tetrahedron 64 (2008) 6475.
[29] G. Oliviero, J. Amato, N. Borbone, S. D'Errico, G. Piccialli, L. Mayol, Tetrahedron Lett. 48 (2007) 397.
[30] G. Oliviero, S. D'Errico, N. Borbone, J. Amato, V. Piccialli, M. Varra, G. Piccialli, L. Mayol, Tetrahedron 66 (2010) 1931.
[31] S. D'Errico, G. Oliviero, J. Amato, N. Borbone, V. Cerullo, A. Hemminki, V. Piccialli, S. Zaccaria, L. Mayol, G. Piccialli, Chem. Commun. 48 (2012) 9310.
[32] G. Oliviero, S. D'Errico, N. Borbone, J. Amato, V. Piccialli, G. Piccialli, L. Mayol, Eur. J. Org. Chem. 2010 (2010) 1517.
[33] S. D'Errico, G. Oliviero, N. Borbone, J. Amato, D. D'Alonzo, V. Piccialli, L. Mayol, G. Piccialli, Molecules 17 (2012) 13036.
[34] D. Musumeci, C. Irace, R. Santamaria, D. Montesarchio, MedChemComm 4 (2013) 1405.
[35] G.N. Roviello, Amino Acids 50 (2018) 933.
[36] N.P. Dolman, J.C.A. More, A. Alt, J.L. Knauss, H.M. Troop, D. Bleakman, G.L. Collingridge, D.E. Jane, J. Med. Chem. 49 (2006) 2579.
[37] G.N. Roviello, A. Ricci, E.M. Bucci, C. Pedone, Mol. BioSyst. 7 (2011) 1773.
[38] C.T. Walsh, W. Zhang, ACS Chem. Biol. 6 (2011) 1000.
[39] S. De, E. Groaz, L. Margamuljana, P. Herdewijn, Chemistry - A European Journal 22 (2016) 8167.
[40] L. Simeone, C. Irace, A. Di Pascale, D. Ciccarelli, G. D'Errico, D. Montesarchio, Eur. J. Med. Chem. 57 (2012) 429.
[41] G.N. Roviello, G. Oliviero, A. Di Napoli, N. Borbone, G. Piccialli, Arabian J. Chem. 13 (2020) 1966.
[42] H. Wang, Z. Feng, Y. Qin, J. Wang, B. Xu, Angew. Chem. Int. Ed. 57 (2018) 4931.
[43] H. Wang, Z. Feng, B. Xu, Theranostics 9 (2019) 3213.
[44] K. Baek, A.D. Noblett, P. Ren, L.J. Suggs, Biomater. Sci. 8 (2020) 3130.
[45] C. Avitabile, C. Diaferia, B. Della Ventura, F.A. Mercurio, M. Leone, V. Roviello, M. Saviano, R. Velotta, G. Morelli, A. Accardo, A. Romanelli, Chem. A Eur. J. 24 (2018) 4729.
[46] C. Avitabile, C. Diaferia, V. Roviello, D. Altamura, C. Giannini, L. Vitagliano, A. Accardo, A. Romanelli, Chem. A Eur. J. 25 (2019) 14850.
[47] C. Coppola, V. Saggiomo, G. Di Fabio, L. De Napoli, D. Montesarchio, J. Organ. Chem. 72 (2007) 9679.
[48] G.N. Roviello, C. Vicidomini, S. Di Gaetano, D. Capasso, D. Musumeci, V. Roviello, RSC Adv. 6 (2016) 14140.
[49] D. Musumeci, A. Mokhir, G.N. Roviello, Bioorg. Chem. 100 (2020) 103862.
[50] D. Musumeci, V. Roviello, G.N. Roviello, Int. J. Nanomed. 13 (2018) 2613.
[51] X. Du, J. Zhou, X. Li, B. Xu, Interface Focus 7 (2017) 20160116.
[52] G.N. Roviello, V. Roviello, I. Autiero, M. Saviano, RSC Adv. 6 (2016) 27607.
[53] M. Pirtskhalava, A. Egoyan, M. Mirtskhulava, G. Roviello, Georgian Med News (2017) 1123.
[54] U. Diederichsen, Angew. Chem., Int. Ed. Engl. 35 (1996) 445.
[55] A.S. Saghyan, H.M. Simonyan, S.G. Petrosyan, A.V. Geolchanyan, G.N. Roviello, D. Musumeci, V. Roviello, Amino Acids 46 (2014) 2325.
[56] A. Carella, V. Roviello, R. Iannitti, R. Palumbo, S. La Manna, D. Marasco, M. Trifuoggi, R. Diana, G.N. Roviello, Int. J. Biol. Macromol. 121 (2019) 77.

[57] M.A. Fik-Jaskółka, A.F. Mkrtchyan, A.S. Saghyan, R. Palumbo, A. Belter, L.A. Hayriyan, H. Simonyan, V. Roviello, G.N. Roviello, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 229 (2020) 117884.
[58] M.A. Fik-Jaskółka, A.F. Mkrtchyan, A.S. Saghyan, R. Palumbo, A. Belter, L.A. Hayriyan, H. Simonyan, V. Roviello, G.N. Roviello, Amino Acids 52 (2020) 755.
[59] C. Di Natale, P.L. Scognamiglio, R. Cascella, C. Cecchi, A. Russo, M. Leone, A. Penco, A. Relini, L. Federici, A. Di Matteo, F. Chiti, L. Vitagliano, D. Marasco, FASEB J. 29 (2015) 3689.
[60] G. Genta-Jouve, L. Weinberg, V. Cocandeau, Y. Maestro, O.P. Thomas, S. Holderith, Chirality 25 (2013) 180.
[61] E. Castiglioni, S. Abbate, F. Lebon, G. Longhi, Chirality 24 (2012) 725.
[62] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J.Y. Xiang, L. Wang, D. Lupyan, M.K. Dahlgren, J.L. Knight, J.W. Kaus, D.S. Cerutti, G. Krilov, W.L. Jorgensen, R. Abel, R.A. Friesner, J. Chem. Theory Comput. 12 (2015) 281.
[63] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, J. Comput. Aided Mol. Des. 27 (2013) 221.
[64] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, J. Med. Chem. 49 (2006) 6177.
[65] K.J. Bowers, D.E. Chow, H. Xu, R.O. Dror, M.P. Eastwood, B.A. Gregersen, J.L. Klepeis, I. Kolossvary, M.A. Moraes, F.D. Sacerdoti, J.K. Salmon, Y. Shan, D.E. Shaw, SC 43 (2006), https://doi.org/10.1109/SC.2006.54.
[66] P. Mark, L. Nilsson, J. Phys. Chem. A 105 (2001) 9954.
[67] C. Sathyaseelan, V. Vijayakumar, T. Rathinavelan, J. Mol. Biol. (2020) 166629.
[68] B.M. Bulheller, J.D. Hirst, Bioinformatics 25 (2009) 539.
[69] F. Menges, Version 1 (2017) 2016.
[70] J.C. Karst, A.C. Sotomayor Pérez, J.I. Guijarro, B. Raynal, A. Chenal, D. Ladant, Biochemistry 49 (2010) 318.
[71] M. Marzano, A.P. Falanga, D. Marasco, N. Borbone, S. D'Errico, G. Piccialli, G.N. Roviello, G. Oliviero, Mar. Drugs 18 (2020) 49.
[72] D.C. Mikles, V. Bhat, B.J. Schuchardt, C.B. McDonald, A. Farooq, Biopolymers 103 (2015) 74.
[73] K.L. Seldeen, C.B. McDonald, B.J. Deegan, V. Bhat, A. Farooq, Biochemistry 48 (2009) 12213.
[74] U. Samanta, D. Pal, P. Chakrabarti, Acta Crystallogr. D Biol. Crystallogr. 55 (1999) 1421.
[75] P.L. Scognamiglio, C. Platella, E. Napolitano, D. Musumeci, G.N. Roviello, Molecules 26 (2021) 3558.
[76] G.N. Roviello, C. Crescenzo, D. Capasso, S. Di Gaetano, S. Franco, E.M. Bucci, C. Pedone, Amino Acids 39 (2010) 795.