



Structure and lowest excitation properties of the backbone-modified GC PNA dimers

K INDUMATHI^a, A ABIRAM^b and G PRAVEENA^{a,*}

^aDepartment of Physics, PSGR Krishnammal College for Women, Coimbatore, India

^bDepartment of Physics, PSG College of Arts and Science, Coimbatore, India

*Corresponding Author. E-mail: praveenag@psgrkcw.ac.in

MS received 28 October 2024; revised 8 April 2025; accepted 23 April 2025

Abstract. The proposed research aims to investigate the structure and lowest excitation properties of GG-CC and GC-CG (G, guanine; C, cytosine) peptide nucleic acid (PNA) dimers, incorporating different amino acids such as serine, aspartic acid, and histidine using the density functional theory (DFT) and time-dependent density functional theory (TDDFT) methods. The structures under consideration have been optimized at the Becke's three-parameter hybrid density functional (B3) with correlation function of Lee, Yang and Parr (LYP)/6-31G* level of theory. The study involves calculating the backbone torsions and backbone-base linker torsions, correlating them with experimental data. The computed excitation energy for the GG-CC PNA system is compared with the natural GG-CC DNA system. The lowest excitation properties, such as excitation energy, wavelength, and oscillator strength, reveal their dependency on both the stacking arrangement and the molecular environment, irrespective of whether the PNA is modified or unmodified. Additionally, a highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) analyses were conducted. This study is intended to serve as a foundational tool for understanding the molecular behaviour of PNAs under light absorption, potentially leading to further exploitation in photo-related applications.

Keywords. Guanine; cytosine; peptide nucleic acid; excitation; density functional theory; time-dependent density functional theory.

1. Introduction

DNA nucleobases are highly photostable due to their ultrafast internal conversion process, which helps in releasing the absorbed excess energy, and thus preventing photodamage in nucleobases from sunlight.¹ Pecourt *et al.*² in their study have demonstrated that the substantial energy of about 34,000 cm⁻¹ gets released by the intramolecular vibrations during the excited state deactivation, as a result of which the thermal equilibrium state of the nucleobases is maintained. The excitation characteristics of DNA molecules explored so far²⁻⁴ depict a vigorous contribution of purine and pyrimidine nucleobases to the main absorption peak around ~260 nm, which corresponds to $n\pi^*$ or $\pi\pi^*$ band transitions. Quite a lot of studies on the excitation relaxation pathways to delve into the intrinsic photophysical and photochemical properties of nucleic acids have been made.⁵⁻⁸ Varsano *et al.*⁹ made an in-depth analysis of the optical absorption spectra of various DNA constructs such as

nucleobases, base pairs, and stacked assemblies using time-dependent density functional theory (TDDFT). Wherein they observed that the change in light absorption strongly depends on the interaction of polarized light with bases, base pairs, π -stacks, or even H-bonds of DNA structures. Sun *et al.*¹⁰ showed the role of base stacking in the shift of the light absorption by studying the lowest excited states of DNA dimers and tetramers using TDDFT.

In addition to the natural DNA molecules, a basic understanding of the photophysical and chemical properties of chemically modified DNA constructs has also evolved to identify a new range of biomolecules, which could be fit into optical and biomedical applications. In this row of research, various nucleic acid analogues such as phosphorothioates, phosphoramidates, methylphosphonates, morpholino phosphorodiamidate, unlocked nucleic acids (UNA), locked nucleic acids (LNAs), threose nucleic acid (TNA), peptide nucleic acids (PNAs)¹¹ and so on are synthesized to attain enhanced properties over natural

nucleic acid systems. For instance, Gazit *et al.*¹² have encountered the fluorescence emission peaks at about 420 to 490 nm in guanine (G) cytosine (C) GC di-PNA (with N-(2-aminoethyl)-glycine) while studying the binding properties of PNA with DNA intercalators. Subsequently, the optoelectronic properties of PNA assemblies as light-emitting diodes were further explored by the group¹³ using the field-effect transistor (FET) device. Furthermore, studies on adopting substituents like fluorenylmethoxycarbonyl (Fmoc) groups and phenylpyrrolocytosine at the backbone of G-enriched-PNA have been explored by various groups for fluorescent properties.^{14–16} Although PNA has found its extraordinary biocompatibility for various applications, there is still room for the development of these chemically modifiable structures.

Nielsen and coworkers^{17,18}, who designed the original PNA, introduced chiral functionalization on PNA backbone to enhance the hybridization and soluble properties with DNA/RNA complexes. In the study, they have shown the discrimination of such chiral functionalization impact on thermal and soluble properties of both PNA/DNA and PNA/RNA complexes relative to their natural counterparts. Subsequently, there are studies which report that using different amino acids instead of the glycine moiety at the PNA backbone has improved efficacy for *in-vivo* biological applications.^{19–21} For instance, Ly *et al.*²² demonstrated the effective in-take of arginine modified guanidinium PNAs by mammalian cells. Studies revealed that using serine instead of glycine at the PNA backbone showed greater solubility.²³ In this line of approach, Shi *et al.*¹⁹ have shown that the inclusion of amino acid residues in the middle of the PNA chain or at N/C termini affords the structural and functional versatility towards the solubility and flexibility of PNA structures. The main reason for the selection of chiral-based PNAs by various groups^{24–26} is due to their position and stereochemical properties that play a crucial role in provoking conformational rearrangement to recognize the specific orientational binding to the target DNA sequences. Alanine and serine incorporated γ -PNAs have been explored by Danith Ly and coworkers^{22,23} wherein their study revealed enhanced hybridization properties of such PNA systems with high sequence selectivity.

Even though quite a lot of studies on the modified and unmodified PNA systems describing their improved properties are available^{27–30}, a study detailing the molecular level intricacies of the structure and photophysical activity of backbone-modified PNA constructs is almost absent in the literature. Moreover, an exploration on the excited state

properties of the PNA constructs with various amino acid incorporation is of paramount importance to infer their performance towards photophysical and bio-related applications. Further, studying insights of the structural and low-lying electronic transition properties of PNAs with G-C would be helpful as it plays a vital role in invading the target nucleotide sequences to form well-ordered PNA: DNA triplexes.³¹ Studying the lowest excited state of GG-CC and GC-CG PNA with amino acid incorporation helps identify how molecular modifications influence peak positions and intensity in absorption/emission spectra. This understanding enables experimentalists to distinguish between fluorescence quenching and spectral shifts, leading to advancements in biosensors, imaging, and material science. Considering the biological importance of the backbone-modified G-containing PNA nanostructural motifs, the present study aims to explore the structural and low-lying electronic transitions of GG-CC and GC-CG PNA dimers with and without modification by applying quantum chemical techniques. Since the polar amino acids are known to be used for their chiral and physiochemical properties in PNA^{19,20}, the present study utilized each one from the three kinds (neutral, basic, and acidic) of polar amino acids, such as serine (Ser^0), aspartic acid (Asp^-), and histidine (His^+) for the PNA backbone modification. The calculated properties, such as excitation energy (eV), wavelength (λ), and oscillator strength (f) will be discussed along with the standard PNA systems. Moreover, such a study on exploring the photophysical properties of the backbone modified PNA dimers at the molecular level is believed to provide deeper insight into the basic structural units of their PNA nanostructures, thereby broadening the perspectives of their contribution in the field of photo-related applications.

2. Systematic illustration and computational details

Of all the studies undertaken so far, the most possible approach which is followed for the modification of PNA is by (i) extending PNA backbone with either of the ethyl to propyl linkage, glycine to the β -alanine linkage, or acetyl to propionyl linkage,³² (ii) adding functional groups at N terminal site of the PNA backbone¹⁴ and (iii) incorporating amino acids at different sites of the PNA backbone²⁶ as shown in Figure 1.

To construct GG-CC and GC-CG dimers with native PNA and modified PNA, a set of GC base pairs

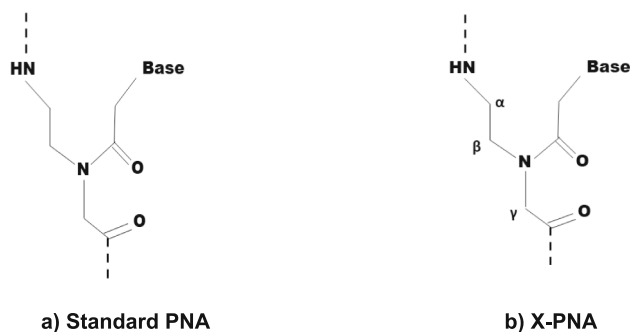


Figure 1. Schematic illustration of (a) standard PNA and (b) X-PNA ($X = \alpha, \beta$ or γ) with the possible sites for modification.

has been stacked at a certain distance and then the backbone of the peptide chain is linked in the specific N sites of the guanine and cytosine. The backbone modifications in GG-CC and GC-CG PNA dimers are made by utilizing different polar amino acids such as serine⁰ (Ser), aspartic acid⁻ (Asp) and histidine⁺ (His), which are positioned at the γ site as depicted in X-PNA (Figure 1). A total of eight dimers have been constructed along with native PNA and labelled as (GG-CC)_G, (GG-CC)_S, (GG-CC)_A, (GG-CC)_H, (GC-CG)_G, (GC-CG)_S, (GC-CG)_A, and (GC-CG)_H throughout the manuscript. Density functional theory (DFT) and time-dependent density functional theory (TDDFT) are popular and efficient methods to study the quantum chemical properties of the ground and excited state structures of various biomolecules. The ground state structures of all eight constructed systems in the gas phase are executed at the B3LYP/6-31G* level of theory using DFT, as shown in Figure 2.

The optimized structures are further subjected to the solution phase (water) based on the integral equation formalism polarizable continuum model (IEFPCM)^{33–35} implicit solvation model with the same level of theory. The backbone torsions such as α' , β' , γ' , δ , ε and ω , and the backbone-base linker torsions, namely χ^1 , χ^2 and χ^3 are calculated for all the modified PNA structures as shown in Figure 3 and are compared with the unmodified PNA results.³⁶

As stated earlier, the immediate response low-lying excited calculations were made for the considered structures in both gas and solution phases using the TDDFT. For a better understanding, the molecular behavioural nature of low-lying excitation, properties include excitation energy (eV), wavelength (λ), and oscillator strength (f) have been calculated. The GAUSSIAN 09W program package³⁷ has been used to carry out all the calculations mentioned above. In addition, we conducted a Frontier molecular orbital analysis (FMOA) to gain insights through the visuals

of orbital distributions upon backbone modification using the GAUSSVIEW 5.0 Program package.³⁸

3. Results and discussion

The structural and low-lying electronic transition state properties of (GG-CC)_G, (GC-CG)_G, (GG-CC)_S, (GG-CC)_A, (GG-CC)_H, (GC-CG)_S, (GC-CG)_A, and (GC-CG)_H have been studied using the B3LYP/6-31G* level of theory. The structural tendencies of all considered systems are analysed through the backbone (α' , β' , γ' , δ , ε and ω) and backbone-base linker (χ^1 , χ^2 and χ^3) torsions. The low-lying electronic transition state properties of these systems are examined through the calculation of excitation energy (eV), wavelength (λ), and oscillator strength (f) for both gas phase and solution phase. The discussion part involves a comparison between standard PNA and backbone modified PNA systems to better understand the PNA activities upon modification.

3.1 Structural analysis of PNA dimeric strands

The backbone (α' , β' , γ' , δ , ε and ω) and backbone-base linker (χ^1 , χ^2 and χ^3) torsion angles for all considered dimeric strands are measured and reported in Table 1 alongside the experimental results of PNA duplexes. The measured backbone and base-linker torsions for the standard PNA system (GG-CC)_G and (GC-CG)_G are almost consistent with the available NMR data³⁹, with notable deviations in β' and ω by $\sim 111.8^\circ$ to 112.5° and 149.5° to 155.2° , respectively.

The torsions α' , γ' and χ^1 for (GG-CC)_G and (GC-CG)_G show opposite sign with slight variations about $\sim 0.4^\circ$, 13.7° and 5.7° and $\sim 7.2^\circ$, 14.8° and 4° when compared to the experimental results. Such deviations observed for the obtained PNA dimeric strands from that of the respective experimental values might be due to the structural constraints considering short PNA dimeric strands rather than PNA duplexes. Calculated torsions α' , β' , γ' , δ , ε , χ^1 , χ^2 and χ^3 for all the modified GG-CC systems are almost comparable with the standard PNA systems, with slight increment in ω by about $\sim 17.6^\circ$ for (GG-CC)_A and $\sim 12.2^\circ$ for (GG-CC)_H. The measured torsions for the hetero stacked systems (GC-CG)_G, (GC-CG)_S, (GC-CG)_A, and (GC-CG)_H indicate that the torsions of (GC-CG)_G and (GC-CG)_S are almost similar. However, a slight discrepancies in the torsions γ' , δ , ε , χ^1 , χ^2 and χ^3 around $\sim 10.8^\circ$, 24.6° , 17.4° , 8.7° , 14.7° , 8.9° for (GC-CG)_A and $\sim 11.1^\circ$, 36.3° , 13.6° , 6.2° , 29.1° , 22.4° for (GC-

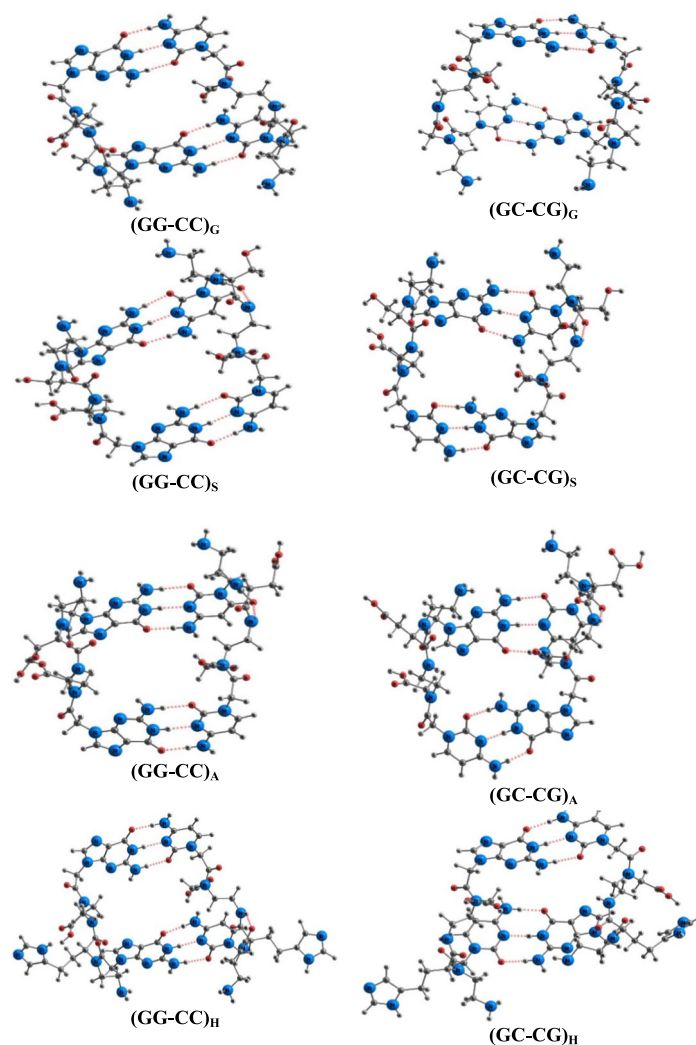


Figure 2. Optimized structures of PNA dimeric strands obtained at B3LYP/6-31G* level of theory.

Table 1. The backbone torsions α' , β' , γ' , δ , ε , ω and backbone base-linker torsions χ^1 , χ^2 and χ^3 ($^\circ$) of all GG-CC and GC-CG PNA dimeric strands obtained at B3LYP/6-31G* level of theory.

PNA dimeric strands	α' ($^\circ$)	β' ($^\circ$)	γ' ($^\circ$)	δ ($^\circ$)	ε ($^\circ$)	ω ($^\circ$)	χ^1 ($^\circ$)	χ^2 ($^\circ$)	χ^3 ($^\circ$)
(GG-CC) _G	93.4 (-93) ^a	-177.8 (-66) ^a	88.7 (-75) ^a	-101.1 (-88) ^a	-134.3 (-155) ^a	25.5 (175) ^a	-7.7 (2) ^a	161.4 (169) ^a	-105.3 (-82) ^a
(GG-CC) _S	103.2	-177.6	89.5	-98.9	-125	21.7	-7.1	155.7	-103.5
(GG-CC) _A	95.7	-176.0	85.6	-113.0	-126.8	43.1	-7.4	153.2	-101.1
(GG-CC) _H	95.9	177.9	93.0	-88.9	-129.2	13.3	-9.0	159.7	-105.4
(GC-CG) _G	100.2 (-93) ^a	-178.5 (-66) ^a	89.8 (-75) ^a	-98.3 (-88) ^a	-125.4 (-155) ^a	22.8 (175) ^a	-6.0 (2) ^a	153.0 (169) ^a	-100.2 (-82) ^a
(GC-CG) _S	99.8	-179.9	89.3	-97.3	-127.5	19.8	-6.6	158.5	-101.7
(GC-CG) _A	109.8	166.3	100.6	-73.7	-108.0	0.8	-14.7	138.3	-109.1
(GC-CG) _H	100.3	158.6	100.9	-62.0	-111.8	-10.2	-12.2	123.9	-122.6

^aThe compared NMR data of backbone and base-linker torsions of PNA-PNA helix are taken from the study of He *et al.*³⁹

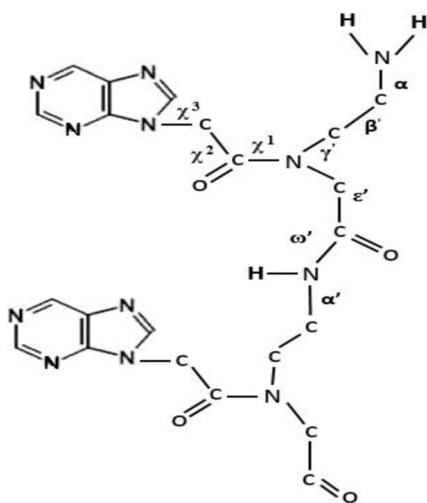


Figure 3. Scheme for torsion and linker angle definitions: $\alpha_{(i+1)} = \text{C-N-C-C}$, $\beta' = \text{N-C-C-N}$, $\gamma' = \text{C-C-N-C}$, $\delta = \text{C-N-C-C}$, $\varepsilon_{(i)} = \text{N-C-C-N}$, $\omega_{(i)} = \text{C-C-N-C}$, $\chi^1 = \text{C-N-C-C}$, $\chi^2 = \text{N-C-C-N1(py) or N9 (pu)}$, $\chi^3 = \text{C-C-N1(py) or N9 (pu)-C2(py) or C4 (pu)}$.

CG)_H with aspartic acid and histidine are noted. The backbone torsion β' of (GC-CG)_A and (GC-CG)_H have opposite signs, with a decrement value of about $\sim 12.2^\circ$ and 19.9° , respectively, compared to (GC-CG)_G. Also, it is noticed that the decrement of ω for (GC-CG)_H is in the range of $\sim 12.6^\circ$ with the opposite sign, whereas for (GC-CG)_A it is decreased by about $\sim 22^\circ$. These fluctuations in torsions might be influenced by the charge of the R groups. The deviations in the torsions of the considered GG-CC and GC-CG PNA systems reveal their structural propensities toward the backbone modification that might be helpful for new designs of highly ordered PNA strands.

3.2 Lowest electronic transition state properties

Studies show that understanding the mechanism of the lowest excitation is pivotal, as it facilitates the exploration of emission properties.⁴⁰ In this study,

Table 2. Singlet electronic transition properties of PNA dimeric strands at B3LYP/6-31G* level of theory in gas phase.

PNA dimeric strands	Transition	Transition of H – L level	Vertical excitation energy (eV)	Wavelength (nm)	Oscillator strength (f)
(GG-CC) _G	S ₁	H-1 to L	3.34	372	0.0011
(GG-CC) _G	S ₂	H to L+1	3.43	361	0.0012
(GG-CC) _G	S ₃	H to L	3.45	359	0
(GG-CC _{DNA}) ^a	G->C; G->C	–	4.83	–	0.01
(GG-CC) _S	S ₁	H-1 to L	3.35	370	0.0007
(GG-CC) _S	S ₂	H to L +1	3.42	362	0.0014
(GG-CC) _S	S ₃	H to L	3.45	360	0
(GG-CC) _A	S ₁	H to L+1	3.43	362	0.0011
(GG-CC) _A	S ₂	H-1 to L	3.43	362	0.0012
(GG-CC) _A	S ₃	H to L	3.47	357	0.0000
(GG-CC) _H	S ₁	H-1 to L	3.39	366	0.0008
(GG-CC) _H	S ₂	H to L	3.42	363	0
(GG-CC) _H	S ₃	H to L+1	3.46	359	0.0011
(GC-CG) _G	S ₁	H-1 to L	3.47	357	0.0006
(GC-CG) _G	S ₂	H to L	3.55	350	0
(GC-CG) _G	S ₃	H to L+1	3.61	344	0.0010
(GC-CG) _S	S ₁	H to L	3.46	358	0.0006
(GC-CG) _S	S ₂	H-1 to L+1	3.59	345	0.0010
(GC-CG) _S	S ₃	H-1 to L	3.61	343	0
(GC-CG) _A	S ₁	H-1 to L	3.50	355	0.0006
(GC-CG) _A	S ₂	H to L	3.52	352	0
(GC-CG) _A	S ₃	H to L+1	3.63	341	0.0010
(GC-CG) _H	S ₁	H to L	3.46	359	0
(GC-CG) _H	S ₂	H-1 to L	3.53	351	0.0008
(GC-CG) _H	S ₃	H to L+1	3.67	338	0.0010

^aThe mentioned excitation energy (average value) and oscillator strength have been calculated by Sun *et al.*¹⁰ for DNA dimers using ω B97X-D/6-31+G (d, p) level of theory.

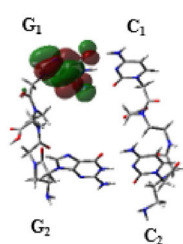
Table 3. Singlet electronic transition state properties of PNA dimeric strands at B3LYP/6-31G* level of theory in solution phase.

PNA dimeric strands	Transition	Transition of H to L level	Vertical excitation energy (eV)	Wavelength (nm)	Oscillator strength (f)
(GG-CC) _G	S ₁	H to L	4.10	302	0.0027
(GG-CC) _G	S ₂	H-1 to L+1	4.17	298	0.0031
(GG-CC) _G	S ₃	H to L+1	4.29	289	0
(GG-CC) _S	S ₁	H to L	4.11	302	0.0019
(GG-CC) _S	S ₂	H-1 to L+1	4.17	297	0.0033
(GG-CC) _S	S ₃	H to L+1	4.31	288	0
(GG-CC) _A	S ₁	H to L	4.11	301	0.0027
(GG-CC) _A	S ₂	H-1 to L+1	4.17	297	0.0032
(GG-CC) _A	S ₃	H to L+1	4.31	287	0
(GG-CC) _H	S ₁	H to L	4.10	302	0.0021
(GG-CC) _H	S ₂	H-1 to L+1	4.15	299	0.0029
(GG-CC) _H	S ₃	H to L+1	4.30	288	0
(GC-CG) _G	S ₁	H to L+1	4.09	303	0.0016
(GC-CG) _G	S ₂	H-1 to L	4.16	298	0.0027
(GC-CG) _G	S ₃	H to L	4.23	293	0.0003
(GC-CG) _S	S ₁	H to L+1	4.08	304	0.0016
(GC-CG) _S	S ₂	H-1 to L	4.17	298	0.0026
(GC-CG) _S	S ₃	H to L	4.22	294	0.0003
(GC-CG) _A	S ₁	H to L	4.11	302	0.0016
(GC-CG) _A	S ₂	H-1 to L+1	4.16	298	0.0025
(GC-CG) _A	S ₃	H to L+1	4.26	291	0.0005
(GC-CG) _H	S ₁	H-1 to L	4.13	300	0.0021
(GC-CG) _H	S ₂	H to L+1	4.16	298	0.0025
(GC-CG) _H	S ₃	H to L	4.23	293	0

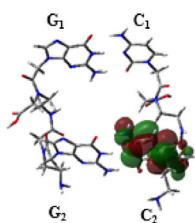
low-lying electronic transitions such as HOMO-1 to LUMO, HOMO to LUMO+1 and HOMO to LUMO were computed for (GG-CC)_G, (GC-CG)_G, (GG-CC)_S, (GG-CC)_A, (GG-CC)_H, (GC-CG)_S, (GC-CG)_A, and (GC-CG)_H. These computations employed the B3LYP/6-31G* level of theory in both gas phase and solution phases, with results presented in Tables 2 and 3, respectively. The electronic transition states identified for all the considered systems fall within the UV spectrum. Specifically, the S₁ transition for (GG-CC)_G, and (GC-CG)_G occurs near 372 nm (3.34 eV) and 357 nm (3.47 eV), respectively. Moreover, the wavelengths (λ) for the S₂ and S₃ transitions of the (GG-CC)_G system are found near 361 nm (3.43 eV) and 359 nm (3.45 eV). For (GC-CG)_G, the λ for S₂ and S₃ transitions fall around 350 nm (3.55 eV) and 344 nm (3.61 eV). Upon examining the oscillator strength for S₁, S₂ and S₃ transitions, it is observed that the S₂ transition of (GG-CC)_G exhibits the maximum absorption with an oscillator strength of ~ 0.0012 f. In contrast, the S₃ transition for (GC-CG)_G shows a maximum absorption of about 0.0010 f. The excitation energy at the maximum absorption for the S₂ transition

in (GG-CC)_G is underestimated by roughly ~ 1.4 eV compared to natural GG-CC_{DNA} systems. Similar trends are observed in systems containing Ser and Asp, where the S₂ transition for (GG-CC)_S occurs at 362 nm with an excitation energy of 3.42 eV, and for (GG-CC)_A 362 nm (3.43 eV). The excitation energy for systems such as (GC-CG) with Ser, Asp, and His increases by ~ 0.07 eV to 0.21 eV compared to that of GG-CC systems. Concurrently, the λ of such systems also decreases by about ~ 15 nm to 25 nm relative to GG-CC systems.

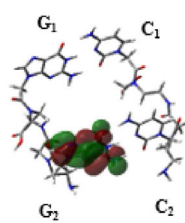
When examining the S₁, S₂, and S₃ transitions in the solution phase, it is noted that the λ values range from ~ 287 nm (4.09 eV) to 304 nm (4.30 eV), which is blue shifted by about ~ 70 nm to 74 nm from that of the gas phase systems. Maximum absorption in the solution phase is observed for (GG-CC)_S and (GC-CG)_G with excitation energies of 4.17 eV (0.0033 f) and 4.16 eV (0.0027 f). The oscillator strength of the considered systems in the solution phase is quite higher of about ~ 0.0010 f and 0.0019 f than the gas phase systems. The quantum changes of excited state properties in both phases underscore the significance of studying the role

(GG-CC)_G

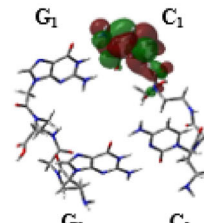
(a) HOMO



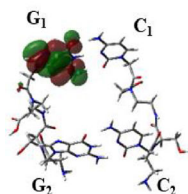
(b) LUMO



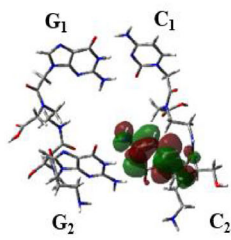
(c) HOMO-1



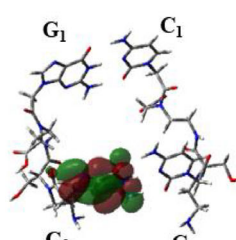
(d) LUMO+1

(GG-CC)_S

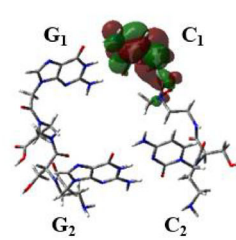
(a) HOMO



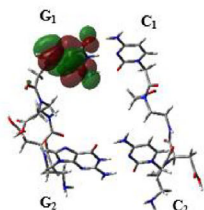
(b) LUMO



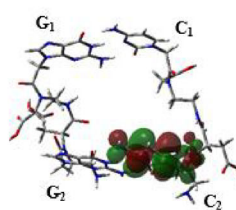
(c) HOMO-1



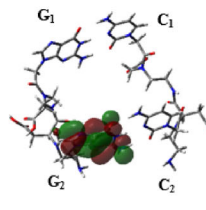
(d) LUMO+1

(GG-CC)_A

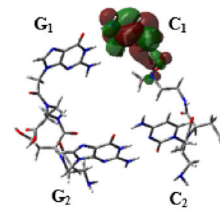
(a) HOMO



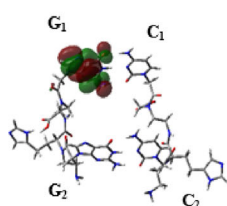
(b) LUMO



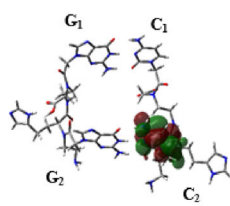
(c) HOMO-1



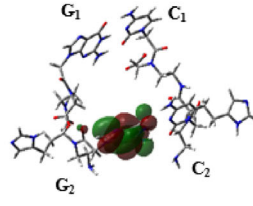
(d) LUMO+1

(GG-CC)_H

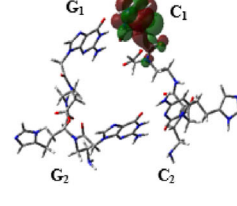
(a) HOMO



(b) LUMO



(c) HOMO-1

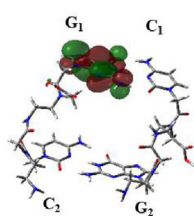


(d) LUMO+1

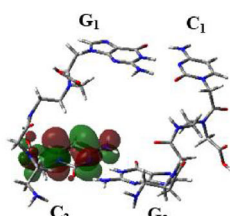
Figure 4. Visual plots of HOMO, LUMO, HOMO-1 and LUMO+1 for GG-CC PNA dimers in gas phase obtained at B3LYP/6-31G* level of theory.

of base stacking towards the photo-related applications. Further, there are studies that are available, which affirm the critical role of base stacking in energy dissipation and the lifetime of excited state.^{10,41} Thus the modifications of the bases in this study demonstrate that

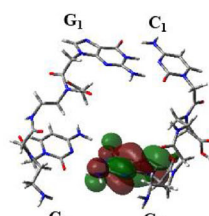
the calculated electronic transition state properties vary in accordance with base stacking configurations. Consequently, we conclude that both the base stacking configuration and the molecular environment significantly impacts excited state properties.

(GC-CG)_G

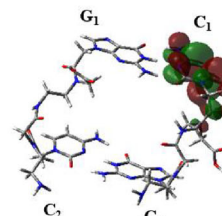
(a) HOMO



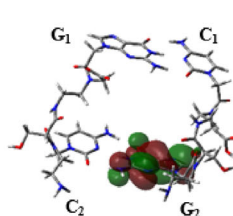
(b) LUMO



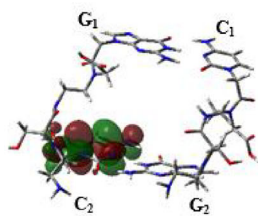
(c) HOMO-1



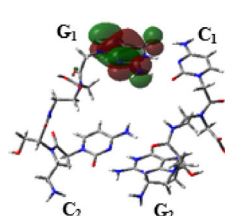
(d) LUMO+1

(GC-CG)_S

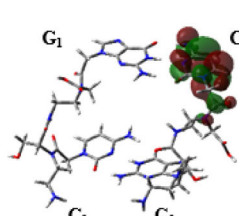
(a) HOMO



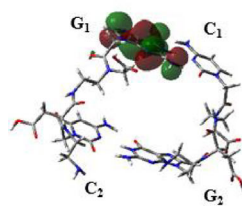
(b) LUMO



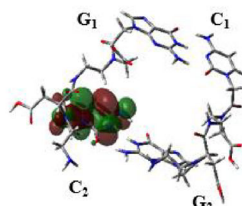
(c) HOMO-1



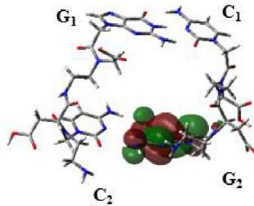
(d) LUMO+1

(GC-CG)_A

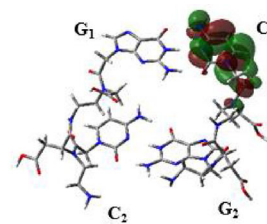
(a) HOMO



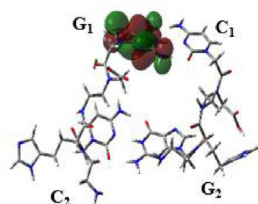
(b) LUMO



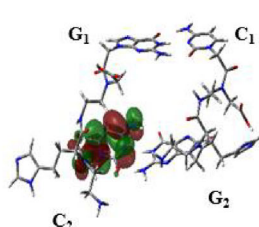
(c) HOMO-1



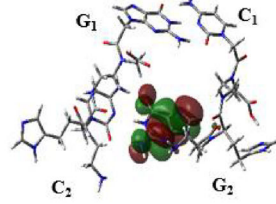
(d) LUMO+1

(GC-CG)_H

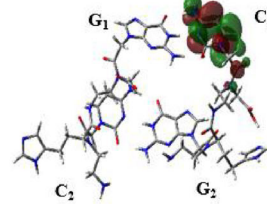
(a) HOMO



(b) LUMO



(c) HOMO-1



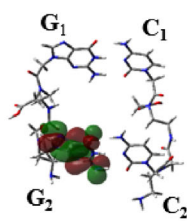
(d) LUMO+1

Figure 5. Visual plots of HOMO, LUMO, HOMO–1 and LUMO+1 for GC-CG PNA dimers in gas phase obtained at B3LYP/6-31G* level of theory.

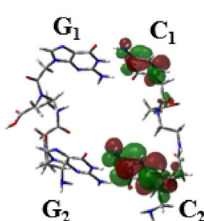
3.3 Orbital distributions in PNA dimeric strands

Typically, the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) are examined to determine the donor and acceptor sites in molecular systems.⁴² In addition, the

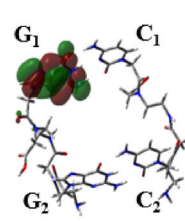
HOMO–1 and LUMO+1 have also been used to understand their contribution in higher excited-state transitions, as it is crucial for charge transfer excitations. While HOMO and LUMO define the primary excitation, these additional orbitals extend the understanding of optical and electronic behaviour,

(GG-CC)_G

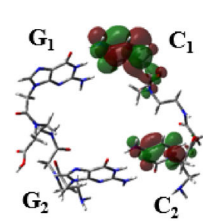
(a) HOMO



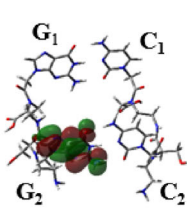
(b) LUMO



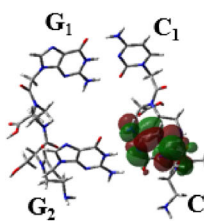
(c) HOMO-1



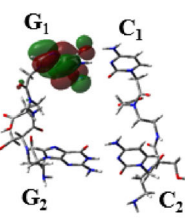
(d) LUMO+1

(GG-CC)_S

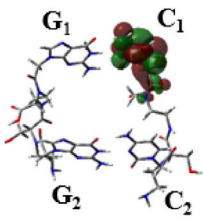
(a) HOMO



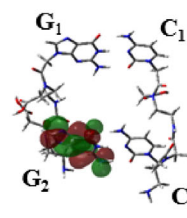
(b) LUMO



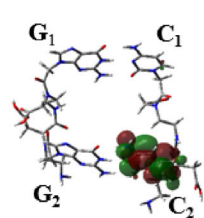
(c) HOMO-1



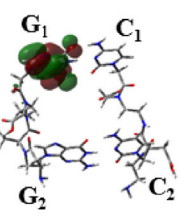
(d) LUMO+1

(GG-CC)_A

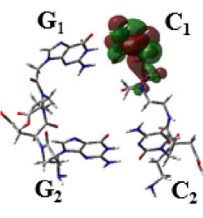
(a) HOMO



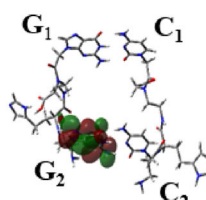
(b) LUMO



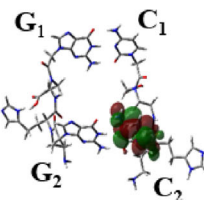
(c) HOMO-1



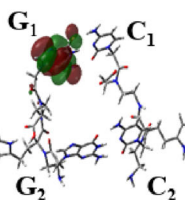
(d) LUMO+1

(GG-CC)_H

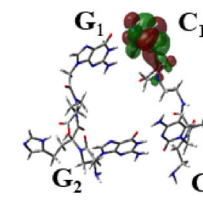
(a) HOMO



(b) LUMO



(c) HOMO-1



(d) LUMO+1

Figure 6. Visual plots of HOMO, LUMO, HOMO-1 and LUMO+1 for GG-CC PNA dimers in solution phase obtained at B3LYP/6-31G* level of theory.

making them essential in computational studies. We have presented in Figures 4, 5, 6 and 7, the orbital plots ranging from HOMO-1 to LUMO (H-1 to L), HOMO to LUMO+1 (H to L+1), and HOMO to LUMO (H to L) for the electronic transitions S_1 , S_2 , and S_3 observed in both gas and solution phases. To enhance the understanding of the orbital localization,

the nucleobases in the stacking arrangements are labelled as G_1 , G_2 , C_1 , and C_2 , as indicated in Figure 4. The charge densities of the HOMO and HOMO-1 in all studied systems (GG-CC)_G, (GG-CC)_S, (GG-CC)_A, (GG-CC)_H, (GC-CG)_G, (GC-CG)_S, (GC-CG)_A, and (GC-CG)_H are primarily located at the purine sites of G_1 and G_2 . Conversely, the LUMO and LUMO+1 are

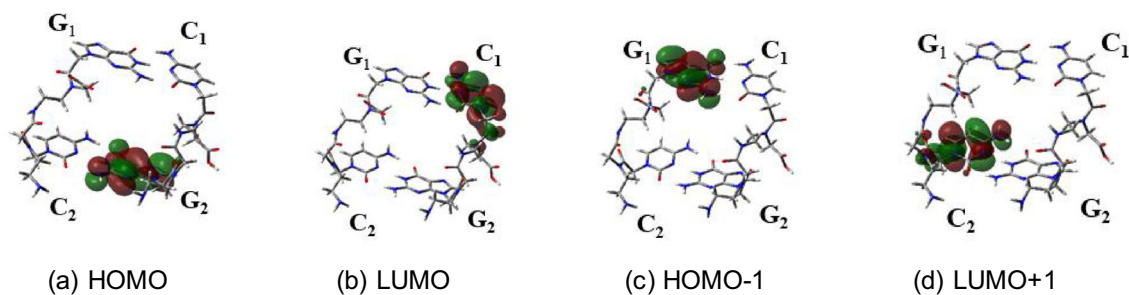
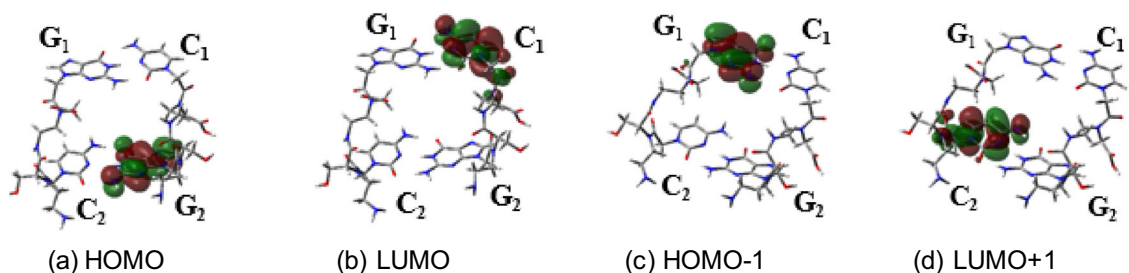
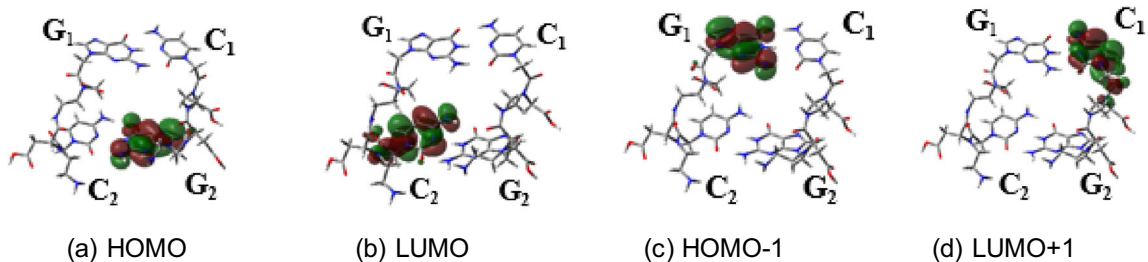
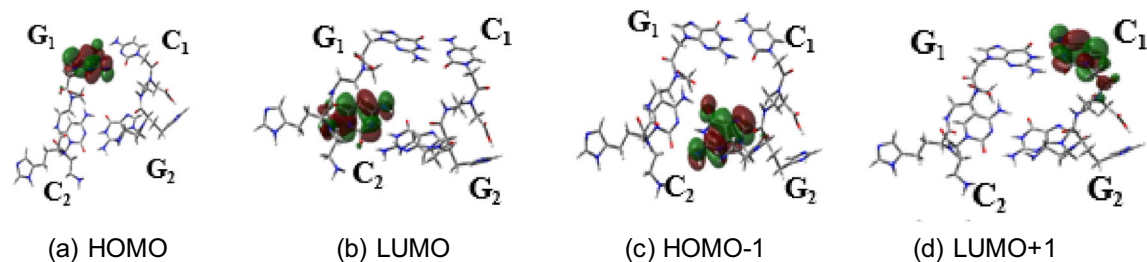
(GC-CG)_G**(GC-CG)_S****(GC-CG)_A****(GC-CG)_H**

Figure 7. Visual plots of HOMO, LUMO, HOMO–1 and LUMO+1 for GC-CG PNA dimers in solution phase obtained at B3LYP/6-31G* level of theory.

predominantly situated at the pyrimidine sites of C₁ and C₂. The HOMO is localized at G₁, and the LUMO+1 is localized at C₁ in the (GG-CC)_G, (GG-CC)_S, and (GG-CC)_H systems, which contributes to the maximum absorption with values of 0.0012 f, 0.0014, and 0.0010 for these systems. A similar transition in GG-CC systems containing Asp occurred

between HOMO-1 and LUMO, which are located at G₂ and C₂, respectively. However, the transition state differed for the (GC-CG)_G, (GC-CG)_A and (GC-CG)_H with that of the (GC-CG)_S, where the transitions from HOMO-1 to LUMO+1 for the (GC-CG)_S and from HOMO to LUMO+1 for the remaining systems also localized at their respective sites of G₁ and C₁.

In the Gly containing GG-CC PNA, it is observed that the HOMO and HOMO-1 are confined at the purine site whereas for the LUMO and LUMO+1, the orbitals are slightly extended towards the peptide chain in the gas phase. The Ser, Asp and His containing GG-CC PNA system also follow the same trend in the orbital localization. This revealed the peptide backbone involvement in the electronic transition of both LUMO and LUMO+1 in the gas phase. The HOMO-1 of Gly, Ser, Asp and His containing GC-CG system, found to have a pearl like an electron orbital distribution over the oxygen in the amide bond (linker of peptide chain) signifies the role of peptide backbone in the charger transfer mechanisms. In the solution phase, the HOMO-1 is present at G₁ and the LUMO+1 at C₁ for the (GG-CC)_S, (GG-CC)_A, and (GG-CC)_H systems with the slight extension of the LUMO and LUMO+1 at the peptide backbone. For the (GG-CC)_G system, HOMO and HOMO-1 is localized at G₁, whereas the LUMO and LUMO+1 is present over both pyrimidines (C₁ and C₂) with the slight extension of the orbitals towards the peptide backbone. The same analysis for GC-CG PNA systems in solution phase reveals that the HOMO and HOMO-1 of Gly and Ser containing systems are located at G₂ and G₁ respectively and the LUMO and LUMO+1 are at C₁ and C₂ is slightly extended towards the peptide backbone. In the case of (GC-CG)_A, the HOMO-1 and LUMO orbital distributions are present at G₁ and C₂ respectively. In the (GC-CG)_H system, the HOMOs are present at G₁ site, and the LUMO+1 is at C₂, with slight extension of orbitals towards the peptide backbone.

4. Conclusion

In this study, we investigated the ground state structure and properties of low-lying excited states of various PNA dimers, including (GG-CC)_G, (GG-CC)_S, (GG-CC)_A, (GG-CC)_H, (GC-CG)_G, (GC-CG)_S, (GC-CG)_A, and (GC-CG)_H, both in gas and solution phases using the B3LYP/6-31G* level of theory. The analysis focussed on the backbone (α , β , γ , δ , ϵ , and ω) and the backbone base-linker (χ^1 , χ^2 , and χ^3) torsions. It was noted that the deviations were significantly greater in the GC-CG PNA systems as compared to the GG-CC PNA systems. Additionally, the lowest excitation analysis indicated that the maximum absorption in the gas phase occurs at 362 nm with an oscillator strength of 0.0014 f for (GG-CC)_S and 345 nm with 0.0010 f for (GC-CG)_S. In contrast, in the solution phase, the maximum absorption for (GG-CC)_S and (GC-CG)_G

was at ~ 297 nm with 0.0033 f and 298 nm with 0.0027 f, respectively. These findings suggest that the wavelength experiences a blue shift when the systems are in a solution environment. Moreover, the oscillator strengths for the GG-CC systems are slightly higher, ~ 0.0001 f to 0.0002 f in the gas phase and ~ 0.0006 f to 0.0016 f in the solution phase, compared to those of the GC-CG systems. The HOMO–LUMO orbital distribution plots provided insights into the exact locations of electron/hole transitions across all the systems studied. In addition, the HOMO-1 and LUMO+1 were also executed. From the extension on the orbital distributions of the LUMO, LUMO+1 and HOMO-1 indicates that the peptide backbone may involve electronic transitions through electronic interaction between the backbone and nucleobases. This results suggest that the modification in the GG-CC and GC-CG PNA backbone have good impact on the electronic transitions further emphasize the role towards the biosensing and optoelectronic applications. Overall, our results conclude that the base stacking configuration significantly influences the structure and excitation properties of the PNA dimeric strands. Furthermore, understanding the impact of the molecular environment on these properties opens new avenues for applications in photo-related technologies.

Acknowledgements

IK is thankful to the GRG Institutions, Coimbatore, India for providing workstation to carry out this work.

References

1. Shukla M K and Leszczynski J 2007 Electronic spectra, excited state structures and interactions of nucleic acid bases and base assemblies: A review *J. Biomol. Struct. Dyn.* **25** 93
2. Pecourt J M, Peon J and Kohler B 2001 DNA Excited-state dynamics: Ultrafast internal conversion and vibrational cooling in a series of nucleosides *J. Am. Chem. Soc.* **123** 10370
3. Schreier W J, Gilch P and Zinth W 2015 Early events of DNA photodamage *Annu. Rev. Phys. Chem.* **66** 497
4. Voet D, Gratzer W B, Cox R A and Doty P 1963 Absorption spectra of nucleotides, polynucleotides, and nucleic acids in the far ultraviolet *Biopolymers* **1** 193
5. Barbatti M, Aquino A J A, Szymczak J J, Nachtigallová D, Hobza P and Lischka H 2010 Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases *Proc. Natl. Acad. Sci. USA* **107** 21453
6. Asturiol D, Lasorne B, Robb M A and Blancafort L 2009 Photophysics of the π , π^* and n, π^* states of thymine: MS-CASPT2 minimum-energy paths and CASSCF on-the-fly dynamics *J. Phys. Chem. A* **113** 10211

7. Sobolewski A L and Domcke W 2004 Ab initio studies on the photophysics of the guanine–cytosine base pair *Phys. Chem. Chem. Phys.* **6** 2763
8. Mennucci B, Toniolo A and Tomasi J 2001 Theoretical study of the photophysics of adenine in solution: tautomerism, deactivation mechanisms, and comparison with the 2 aminopurine fluorescent isomer *J. Phys. Chem. A* **105** 4749
9. Varsano D, Di Felice R, Marques M A and Rubio A A 2006 TDDFT Study of the excited states of DNA bases and their assemblies *J. Phys. Chem. B* **110** 7129
10. Sun H, Zhang S, Zhong C and Sun Z 2016 Theoretical study of excited states of DNA base dimers and tetramers using optimally tuned range-separated density functional theory *J. Comput. Chem.* **37** 684
11. Wang F, Li P, Chu H C and Lo P K 2022 Nucleic acids and their analogues for biomedical applications *Biosensors* **12** 93
12. Berger O, Adler-Abramovich L, Levy-Sakin M, Grunwald A, Liebes Y, Bachar M *et al.* 2015 Light-emitting self-assembled peptide nucleic acids exhibit both stacking interactions and Watson–Crick base pairing *Nat. Nanotechnol.* **10** 353
13. Berger O and Gazit E 2017 Molecular self-assembly using peptide nucleic acids. *Biopolymers* **108**
14. Avitabile C, Diaferia C, Della Ventura B, Mercurio F A, Leone M, Roviello V *et al.* 2018 Self-assembling of Fmoc-GC peptide nucleic acid dimers into highly fluorescent aggregates *Chem. Eur. J.* **24** 4729
15. Wojciechowski F and Hudson R H 2008 Fluorescence and hybridization properties of peptide nucleic acid containing a substituted phenylpyrrolocytosine designed to engage guanine with an additional H-bond *J. Am. Chem. Soc.* **130** 12574
16. Basavalingappa V, Bera S, Xue B, Azuri I, Tang Y, Tao K *et al.* 2019 Mechanically rigid supramolecular assemblies formed from an Fmoc-guanine conjugated peptide nucleic acid *Nat. Commun.* **10** 5256
17. Dueholm K L, Petersen K H, Jensen D K, Egholm M, Nielsen P E and Buchardt O 1994 Peptide nucleic acid (PNA) with a chiral backbone based on alanine *Bioorg. Med. Chem. Lett.* **4** 1077
18. Nielsen P E, Haaima G, Lohse A and Buchardt O 1996 Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: Hybridization and solubility properties of D-lysine PNA *Angew. Chem. Int. Ed. Engl.* **35** 1939
19. Duan T, He L, Tokura Y, Liu X, Wu Y and Shi Z 2018 Construction of tunable peptide nucleic acid junctions *Chem. Comm.* **54** 2846
20. Ryan C A and Rozners E 2020 Impact of chirality and position of lysine conjugation in triplex-forming peptide nucleic acids *ACS Omega* **5** 28722
21. Wu Y and Xu J C 2001 Synthesis of chiral peptide nucleic acids using Fmoc chemistry *Tetrahedron* **57** 8107
22. Zhou P, Dragulescu-Andrasi A, Bhattacharya B, O'Keefe H, Vatta P, Hyldig-Nielsen J J and Ly D H 2006 Synthesis of cell-permeable peptide nucleic acids and characterization of their hybridization and uptake properties *Bioorg. Med. Chem. Lett.* **16** 4931
23. Dragulescu-Andrasi A, Rapireddy S, Frezza B M, Gayathri C, Gil R R and Ly D H 2006 A simple gamma-backbone modification preorganizes peptide nucleic acid into a helical structure *J. Am. Chem. Soc.* **128** 10258
24. Suparpprom C and Vilaivan T 2022 Perspectives on conformationally constrained peptide nucleic acid (PNA): insights into the structural design, properties and applications *RSC. Chem. Biol.* **3** 648
25. Sugiyama T and Kittaka A 2012 Chiral peptide nucleic acids with a substituent in the N-(2-aminoethoxy)glycine backbone *Molecules* **18** 287
26. Moccia M, Adamo M F A and Saviano M 2014 Insights on chiral, backbone modified peptide nucleic acids: properties and biological activity *Artif. DNA: PNA XNA* **5** e1107176
27. Englund E A and Appella D H 2005 Synthesis of gamma-substituted peptide nucleic acids: a new place to attach fluorophores without affecting DNA binding *Org. Lett.* **7** 3465
28. De Koning M C, Petersen L, Weterings J J, Overhand M, van der Marel G A and Filippov D V 2006 Synthesis of thiol-modified peptide nucleic acids designed for post-assembly conjugation reactions *Tetrahedron* **62** 3248
29. Sforza S, Tedeschi T, Corradini R and Marchelli R 2007 Induction of helical handedness and DNA binding properties of peptide nucleic acids (PNAs) with two stereogenic centres *Eur. J. Org. Chem.* **2007** 5879
30. Gupta A, Mishra A and Puri N 2017 Peptide nucleic acids: Advanced tools for biomedical applications *J. Biotechnol.* **259** 148
31. Panyutin I G, Onyshchenko M I, Englund E A, Appella D H and Neumann R D 2012 Targeting DNA G-quadruplex structures with peptide nucleic acids *Curr. Pharm. Des.* **18** 1984
32. Hyrup B, Egholm M, Nielsen P E, Wittung P, Norden B and Buchardt O 1994 Structure activity studies of the binding of modified peptide nucleic-acids (Pnas) to DNA *J. Am. Chem. Soc.* **116** 7964
33. Cancès E, Mennucci B and Tomasi J 1997 A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics *J. Chem. Phys.* **107** 3032
34. Mennucci B, Cancès E and Tomasi J 1997 Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: Theoretical bases, computational implementation, and numerical applications *J. Phys. Chem. B* **101** 10506
35. Rega N, Cossi M and Barone V 1999 Improving performance of polarizable continuum model for study of large molecules in solution *J. Comput. Chem.* **20** 1186
36. Topham C M and Smith J C 1999 The influence of helix morphology on co-operative polyamide backbone conformational flexibility in peptide nucleic acid complexes *J. Mol. Biol.* **292** 1017
37. Frisch M J, Trucks G W, Schlegel H B, Scuseria G E, Robb M A, Cheeseman J R *et al.* (Gaussian 09, Revision A.02)
38. Dennington R and Keith J M T 2009 Semichem Inc.: Shawnee Mission KS.

39. He W, Hatcher E, Balaeff A, Beratan D N, Gil R R, Madrid M and Achim C 2008 Solution structure of a peptide nucleic acid duplex from NMR data: features and limitations *J. Am. Chem. Soc.* **130** 13264
40. Reveguk Z V, Khoroshilov E V, Sharkov A V, Pomogaev V A, Buglak A A, Tarnovsky A N and Kononov A I 2019 Exciton absorption and luminescence in i-motif DNA *Sci. Rep.* **9** 15988
41. Schimelman J B, Dryden D M, Poudel L, Krawiec K E, Ma Y, Podgornik R *et al.* 2015 Optical properties and electronic transitions of DNA oligonucleotides as a function of composition and stacking sequence *Phys. Chem. Chem. Phys.* **17** 4589
42. Labet V, Morell C, Tognetti V, Syzgantseva O, Joubert L, Grand A and Cadet J 2014 Characterization of the chemical reactivity and selectivity of DNA bases through the use of DFT-based descriptors structure, bonding and reactivity of heterocyclic compounds *Top. Heterocycl. Chem.* **38** 35

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.