


Article

Nucleobase–Guanidiniocarbonyl-Pyrrole Conjugates as Novel Fluorimetric Sensors for Single Stranded RNA

Željka Ban ¹, Biserka Žinić ¹, Robert Vianello ¹ , Carsten Schmuck ² and Ivo Piantanida ^{1,*}

¹ Division of Organic Chemistry and Biochemistry, Ruder Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia; Zeljka.Ban@irb.hr (Ž.B.); biserka.zinic@irb.hr (B.Ž.); robert.vianello@irb.hr (R.V.)

² Institute for Organic Chemistry, University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany; carsten.schmuck@uni-due.de

* Correspondence: Ivo.Piantanida@irb.hr; Tel.: +385-1-4571-326

Received: 1 November 2017; Accepted: 9 December 2017; Published: 13 December 2017

Abstract: We demonstrate here for the first time that a guanidiniocarbonyl-pyrrole (GCP) unit can be applied for the fine recognition of single stranded RNA sequences—an intuitively unexpected result since so far binding of the GCP unit to ds-DNA or ds-RNA relied strongly on minor or major groove interactions, as shown in previous work. Two novel nucleobase–GCP isosteric conjugates differing in the flexibility of GCP unit revealed a fluorimetric recognition of various single stranded RNA, which could be additionally regulated by pH. The more rigid conjugate showed a specific fluorescence increase for poly A only at pH 7, whereby this response could be reversibly switched-off at pH 5. The more flexible derivative revealed selective fluorescence quenching by poly G at pH 7 but no change for poly A, whereas its recognition of poly AH⁺ can be switched-on at pH 5. The computational analysis confirmed the important role of the GCP fragment and its protonation states in the sensing of polynucleotides and revealed that it is affected by the intrinsic dynamical features of conjugates themselves. Both conjugates showed a negligible response to uracil and cytosine ss-RNA as well as ds-RNA at pH 7, and only weak interactions with ds-DNA. Thus, nucleobase–GCP conjugates can be considered as novel lead compounds for the design of ss-RNA or ss-DNA selective fluorimetric probes.

Keywords: pyrrolocytosine; guanidiniocarbonyl-pyrrole; DNA/RNA binding; fluorimetric recognition

1. Introduction

Molecular visualization and recognition of DNA and RNA are an important area of ongoing research in medicine and molecular biology, in which small molecules can be not only useful dyes but also effective anticancer, antibiotic and antiviral therapeutic agents [1]. Besides the naturally most abundant ds-DNA, many DNA and RNA structures were discovered (i-motif, triplex and G-quadruplex structures), becoming novel targets for drugs, for instance in cancer chemotherapy [2]. Most of the studies of small molecules were focused on their interactions with double- or multiple-stranded DNA or RNA, but not with single stranded DNA/RNA. That was mostly due to a lack of interest since small molecules generally bind more strongly to multi-stranded nucleic acid structures than to single-stranded (ss) polymers. Moreover, ss-polynucleotides adopt poorly defined secondary structures, making ligand design a very challenging task. Nevertheless, single-stranded nucleic acids such as poly(A) have critical roles in cell biology [3–5], while single stranded G-rich structures can fold in a variety of G-quadruplexes. Furthermore, some RNA in the genomes of various cardioviruses and encephalomyocarditis viruses contain stretches composed of more than 75% cytosine, and single stranded C-polymers can adopt a non-B form structure called i-motif under slightly acidic or neutral pH [6,7]. Very few small molecules reveal selective binding to single stranded DNA or RNA structures.

For instance, some isoquinoline alkaloids bind poly(A) with considerable affinity [8,9] as does the aminoglycoside antibiotic neomycin [10]. The most effective approach to ss-polynucleotide recognition by small molecules (MW < 1000) was combining intercalative unit with a nucleobase, as introduced by Lhomme et al. [11,12], and developed for various ss-polynucleotides by a series of phenanthridinium [13–17] and bis-phenanthridinium-nucleobase conjugates [18,19]. Also, some DBTAA-adenine conjugates specifically recognised poly dT [20].

Our research on novel moieties for interactions with DNA/RNA revealed intriguing properties of the guanidiniocarbonyl-pyrrole system (abbreviated GCP), which we studied in detail on various ds-DNA and ds-RNA. Particularly interesting were aryl-GCP conjugates, which, depending on their structure, showed specific fluorimetric and CD recognition of different ds-DNA or ds-RNA structures depending on pH or steric and/or H-bonding properties of conjugates [21–25]. Mostly, the GCP unit interacted within the DNA or RNA grooves, although there were indications that in its protonated form (pH 5) GCP could interact also with negatively charged DNA/RNA backbone. Particularly latter property raised the question whether GCP unit could also be used for interactions with DNA/RNA targets which do not have a well-defined secondary structure and hence minor or major grooves, respectively.

Up till now, we did not assay the interactions of aryl-GCP conjugates with single stranded DNA or RNA. To address this, we relied on our previous research based on condensed-aryl-nucleobase conjugates [13–20] targeting particular ss-polynucleotide sequences, whereby large aryl moiety ensured sufficient DNA/RNA affinity and spectroscopic response while an attached nucleobase contributed to particular ss-sequence recognition by a set of H-bonds. In the research presented here we designed and prepared a novel fluorescent cytosine-amino acid analogue **1** with C- and N-end available for peptide coupling to ensure broad applicability (Figure 1). Further on, we attached GCP on lysine at two different positions and consequently coupled the two different GCP-lysines with **1**. The obtained nucleobase–GCP conjugates are isosteres, differing in GCP position either on the shorter and more rigid alpha-amino position (**2**) or on the longer, more flexible side-chain amino position (**3**).

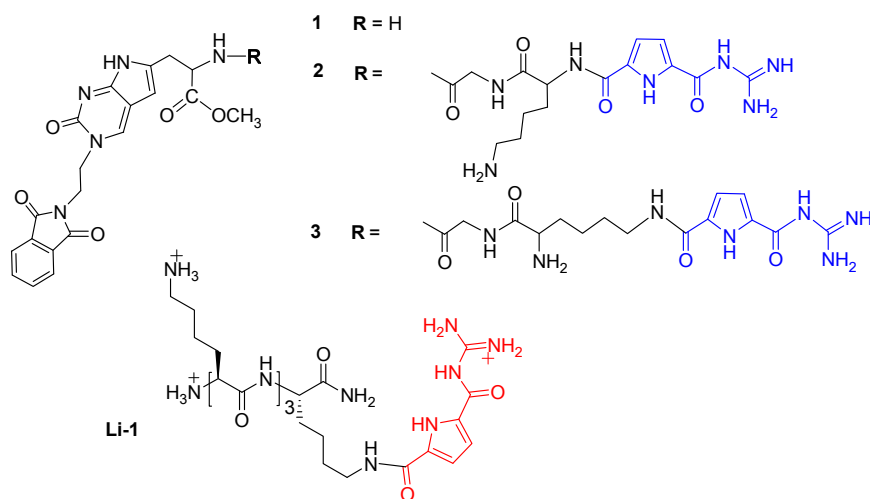


Figure 1. Studied compounds **1**–**3** and previously studied analogue **Li-1** [26].

2. Results and Discussion

2.1. Chemistry

We have designed and synthesized a new type of molecular probes by combining two molecular fragments Pyrr-C and GCP, both well known for their capability to strongly interact with nucleic acids. The fluorescent pyrrolocytosine nucleobase (Pyrr-C) [27] preserves the H-bonding pattern of the parent

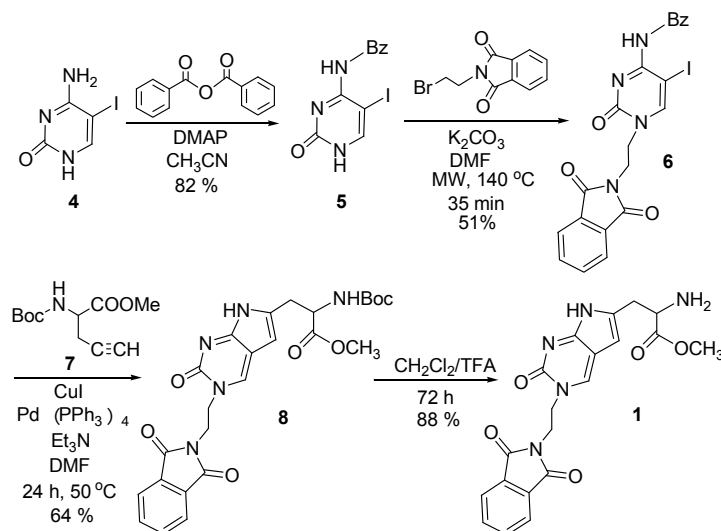
nucleobase cytosine [28,29], while the guanidiniocarbonyl-pyrrole (GCP) moiety is known as a highly efficient DNA minor groove binder [23,25].

We synthesized the fluorescent Pyrr-C-amino acid analogue **1** (Scheme 1) which served as the starting compound for the preparation of the two novel nucleobase-guanidiniocarbonyl-pyrrole conjugates **2** and **3**. Previous studies have shown that the Sonogashira coupling of terminal alkynes with 5-iodocytosine derivatives gave the 5-alkynyl products while *N*⁴-benzoyl-5-iodocytosine derivatives lead to the annulated pyrrolocytosine derivatives via domino cross-coupling and cyclization reaction [30,31].

With these facts in mind, the 5-iodo-*N*⁴-benzoylcytosine **5** was synthesized in a very good yield starting from cytosine (which was previously iodinated by using I₂/HIO₃ in biphasic AcOH/H₂O/CH₂Cl₂ yielding 5-iodocytosine **4** [32,33], followed by the reaction with benzoic anhydride and the presence of DMAP in acetonitrile as described earlier for the synthesis of *N*⁴-benzoylcytosine (Scheme 1) [34].

As the next step, we planned to introduce N-1 substituent by alkylation of **5** with *N*-(2-bromoethyl)phthalimide. This substituent may have multiple roles; (i) it may simplify the isolation of the subsequent products due to increased lipophilicity; (ii) after removal of the phthalimide group a free amino group is generated enabling further functionalization; and (iii) the phthalimide groups itself could contribute to intermolecular interaction by H-bonding and/or aromatic stacking interactions. The best yield of N-1 alkylated product was obtained by alkylation of **5** with *N*-(2-bromoethyl)phthalimide in the presence of K₂CO₃/DMF under microwave irradiation at 140 °C giving **6** in 51% yield.

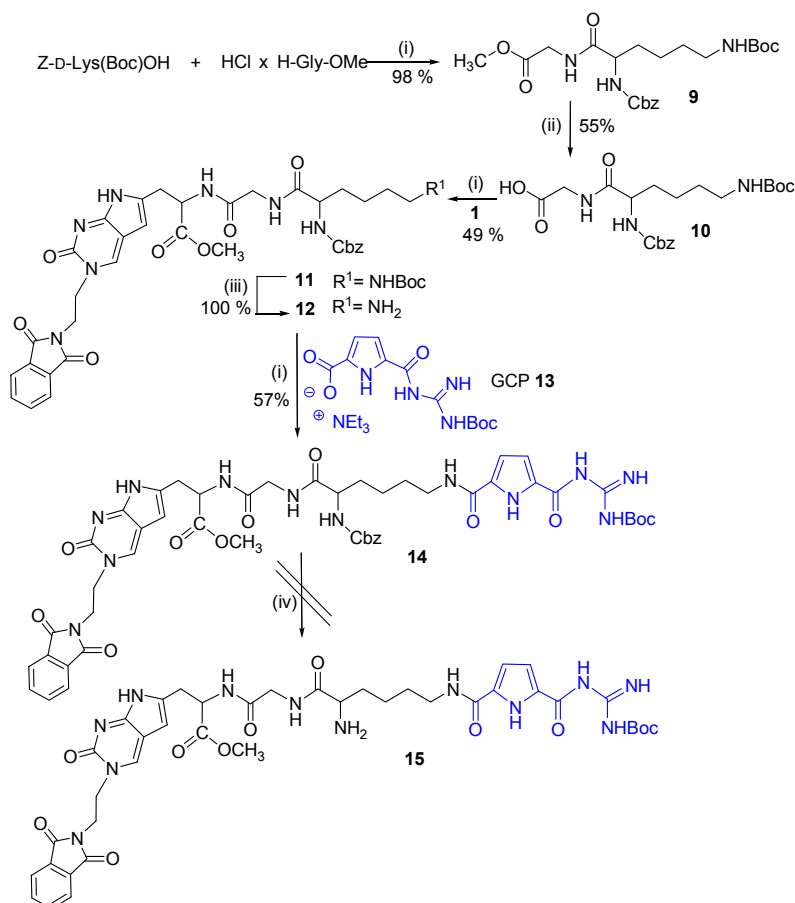
Pyrr-C derivative **8** was synthesized by adopting the methods reported in the literature [35]. The one pot sequential Sonogashira cross-coupling reaction and annulation of **6** with protected [36] propargylglycine **7** was used. Compound **8** was readily deprotected by CH₂Cl₂/TFA to afford **1** in a very good yield (88%).



Scheme 1. Synthesis of novel fluorescent Pyrr-C-amino acid analogue **1**.

Having the fluorescent Pyrr-C-amino acid analogue **1** in hand, we decided to link it with the guanidiniocarbonyl-pyrrole unit (Boc-GCP-OH) [37] **13** through a Gly-Lys dipeptide bridge (Scheme 2). We planned to first attach the fluorescent Pyrr-C unit **1** to Gly-Lys dipeptide and the GCP **13** unit in the second step (Scheme 2). Starting from the commercially available Z-D-Lys(Boc)-OH and H-Gly-OMe hydrochloride, the Z-D-Lys(Boc)-Gly-OMe dipeptide **9** was synthesized in excellent yield by a standard coupling reaction using HBTU/HOBt as coupling reagents and the base, triethylamine in acetonitrile [38]. Dipeptide **9** was readily deprotected with 2 M NaOH in dioxane/H₂O (*v/v* 3:1)

giving **10**. The same coupling procedure was applied for coupling of the Pyrr-C amino acid analogue **1** with Z-D-Lys(Boc)-Gly-OH dipeptide **10**; the product **11** is obtained in moderate yield. The Boc-protecting group was removed with a mixture of CH₂Cl₂/TFA affording **12** in quantitative yield. In the next step, **12** was coupled with GCP **13** giving the conjugate **14**. Unexpectedly, the removal of Cbz protecting group in **14** by hydrogenolysis failed and no deprotected compound **15** could be isolated.

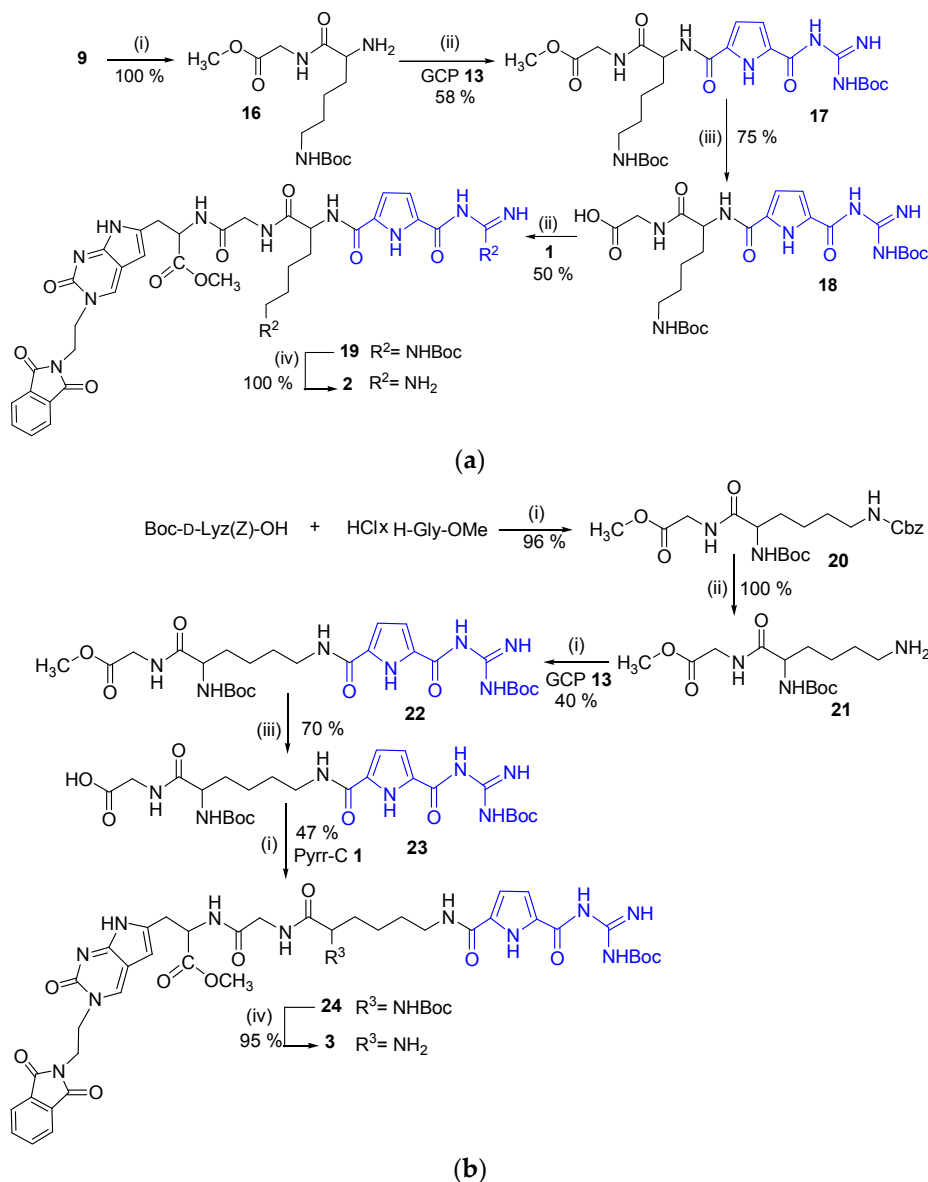


Scheme 2. Attempted synthesis of **15**. *Reagents and conditions:* (i) HBTU, HOBt, Et₃N, CH₃CN; (ii) 2 M NaOH, dioxane/water (*v/v* 3:1); (iii) TFA/CH₂Cl₂; (iv) H₂, Pd/C, MeOH.

Hence, the synthetic approach was changed in the way that the GCP **13** is firstly attached to Gly-Lys dipeptide at either of the two (alpha-amino or side-chain amino) positions, and the obtained products are subsequently coupled with the fluorescent Pyrr-C unit **1** (Scheme 3a,b).

The synthesis of novel nucleobase-guanidiniocarbonyl-pyrrole conjugate **2** started by hydrogenolysis of Z-D-Lys(Boc)-Gly-OMe dipeptide **9** which afforded MeO-Gly-H-Lys(Boc)-dipeptide **16** in quantitative yield (Scheme 3a). Using a standard coupling reaction with HBTU/HOBt conjugate **17** was synthesized in moderate yield. The methyl ester was then removed by basic hydrolysis yielding **18**, which was then coupled with Pyrr-C derivative **1** resulting in hybrid compound **19**. Finally desired conjugate **2**, with the GCP-unit attached to the α -amino position of Lys unit, was obtained by deprotection of Boc-protecting groups with CH₂Cl₂/TFA. The preparation of nucleobase-GCP conjugate **3** with the GCP-unit attached to the side-chain amino position of Lys unit is outlined in Scheme 3b. In this case, the starting materials were amino acid Boc-D-Lys(Z)-OH and H-Gly-OMe hydrochloride, which were coupled by HBTU/HOBt activation giving the dipeptide **20**. After the Cbz deprotection, **21** was obtained in quantitative yield. Then, **21** was coupled with guanidiniocarbonyl-pyrrole derivative **13** under standard HBTU/HOBt conditions, providing **22**

in moderate yield. The methyl ester in the conjugate **22** was easily removed by basic hydrolysis giving the GCP peptide **23** which was then coupled with the Pyrr-C derivative **1** under standard coupling conditions giving the conjugate **24**. Finally, Boc removal from **24** was again carried out with TFA, providing the Pyrr-C-guanidiniocarbonyl-pyrrole derivative **3** in high yield.



Scheme 3. Synthesis of nucleobase-GCP conjugates **2** and **3**. *Reagents and conditions:* (a) (i) H₂, Pd/C, MeOH; (ii) HBTU, HOBT, Et₃N, CH₃CN; (iii) LiOH, THF/water (v/v 4:1); (iv) CH₂Cl₂/TFA; (b) (i) HBTU, HOBT, Et₃N, CH₃CN; (ii) H₂, Pd/C, MeOH; (iii) LiOH, THF/water (v/v 4:1); (iv) TFA/CH₂Cl₂.

2.2. Interactions with DNA and RNA

All compounds **1–3** are moderately soluble in water, and for easier application 0.01 M stock solutions were prepared in DMSO and diluted in buffered solutions prior to use (0.05 M Na cacodylate, DMSO content of the final solutions <0.01%). The UV/Vis spectra (Figure 2) of buffered solutions of **1–3** are proportional to their concentrations up to $c = 2 \times 10^{-5}$ mol dm^{−3}, pointing out that studied compounds do not aggregate by intermolecular stacking at experimental conditions used. Since it is known that GCP unit is protonated at weakly acidic conditions (pH 5) [25], UV/Vis spectra were

collected at pH 7 and pH 5, but only negligible differences were observed. At pH 7 compounds **1–3** bear one positive charge, while at pH 5 the GCP group is also protonated [25], so **2, 3** have charge +2. Absorption maxima and corresponding molar extinction coefficients (ϵ) are given in Table S1 (Supplementary Materials). Heating of the aqueous solutions of **1–3** up to 90 °C did not cause any significant changes in the UV/Vis spectra and reproducibility upon cooling back to room temperature verified the chemical stability of the compounds.

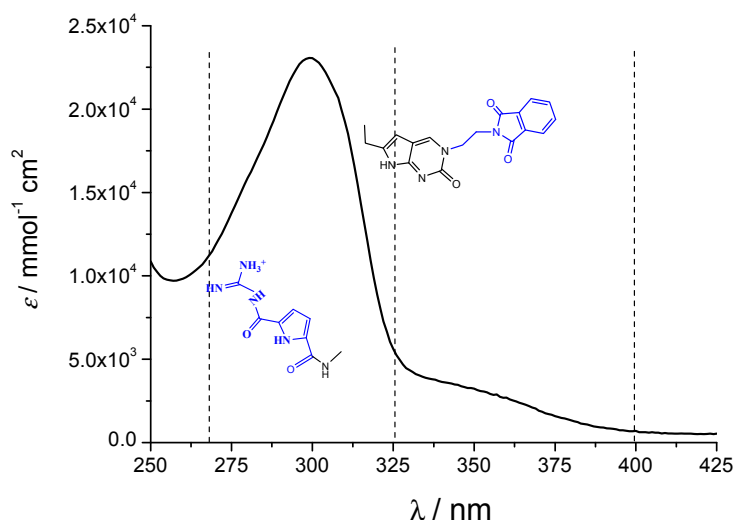


Figure 2. UV/Vis spectrum of **2** at pH 7.0, sodium cacodylate buffer, $I = 0.05$ M, note marked absorption ranges for two chromophores.

The studied compounds have two fluorophores (GCP and nucleobase) connected by a more (compound **3**) or less (compound **2**) flexible linker. As shown in Figures 2 and 3 (and Figure S12 of Supplementary Materials for reference compound **1**), excitation at $\lambda_{\text{exc}} = 299$ nm excites all fluorophores, while excitation at $\lambda_{\text{exc}} = 355$ nm excites only the nucleobase. That allows sequential fluorimetric studies, whereby fluorimetric response upon excitation at $\lambda_{\text{exc}} = 299$ nm monitors the impact of studied binding to all chromophores, while excitation at $\lambda_{\text{exc}} = 355$ nm is focused on the changes induced only on nucleobase-phthalimide moiety.

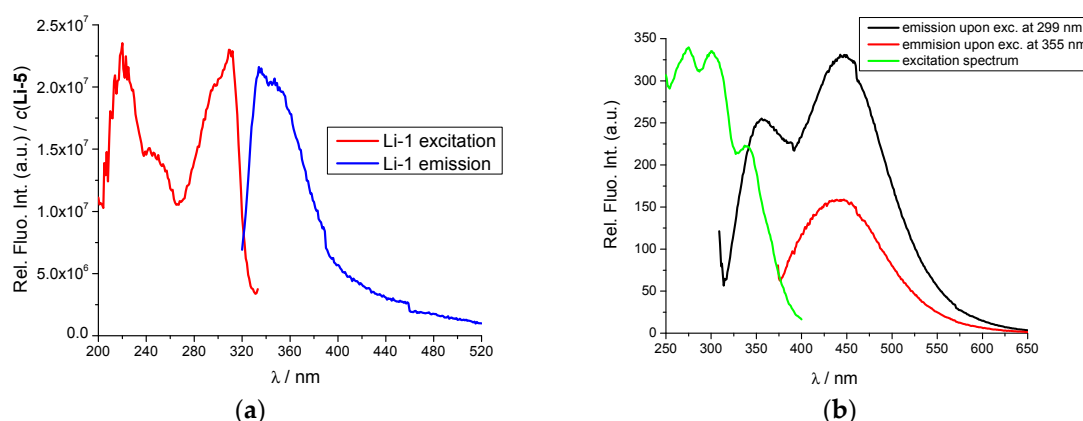


Figure 3. (a) Previously reported analogue GCP [26] fluorescence excitation spectrum ($\lambda_{\text{em}} = 350$ nm) and emission spectrum ($\lambda_{\text{exc}} = 299$ nm); (b) excitation ($\lambda_{\text{em}} = 447$ nm) and emission ($\lambda_{\text{exc}} = 299$ nm, $\lambda_{\text{exc}} = 355$ nm) spectra of **2**. Done at pH 7.0, sodium cacodylate buffer, $I = 0.05$ M.

2.3. DNA/RNA Binding Studies

Interactions of **1–3** with DNA and RNA were studied at both pH 7 and pH 5, due to the difference in the protonation state of GCP [25]. It should be stressed that poly A and poly C are single stranded under neutral conditions (pH 7), while at pH 5 they become double stranded (poly AH⁺-poly AH⁺ and poly CH⁺-poly CH⁺) [19,39].

With representative ds-DNA (calf thymus DNA) as well as with ds-RNA (poly A-poly U) none of the studied compounds showed any stabilisation in thermal denaturation experiments, at both pH 7 and pH 5 (Figures S46 and S47 of Supplementary Materials). Lack of thermal stabilisation excludes intercalative binding mode [2].

Fluorimetric titrations of **2** (Figure 4) and **3** (Figure S29 of Supplementary Materials) at pH 7 revealed significant emission increase only upon addition of ct-DNA, but no emission change upon titration with ds-RNA (poly A-poly U, Figures S19 and S33 of Supplementary Materials). Intriguingly, at pH 5 also poly A-poly U yielded significant increase of fluorescence (Figure S24 of Supplementary Materials), but only for **2**. The lack of binding of **2, 3** to ds-RNA at pH 7, but efficient binding of **2** at pH 5, could be partially attributed to the protonation of the GCP unit, which as previously shown, can significantly contribute to polynucleotide binding [22–25]. However, obviously position of GCP unit in the compound is also important for efficient binding, thus only for **2** but not **3** giving measurable fluorimetric changes are observed. Processing of titration data by Scatchard equation [40,41] revealed similar affinities ($\log K_s \approx 4$, Table 1) for all ds-DNA and ds-RNA (RNA only at pH 5).

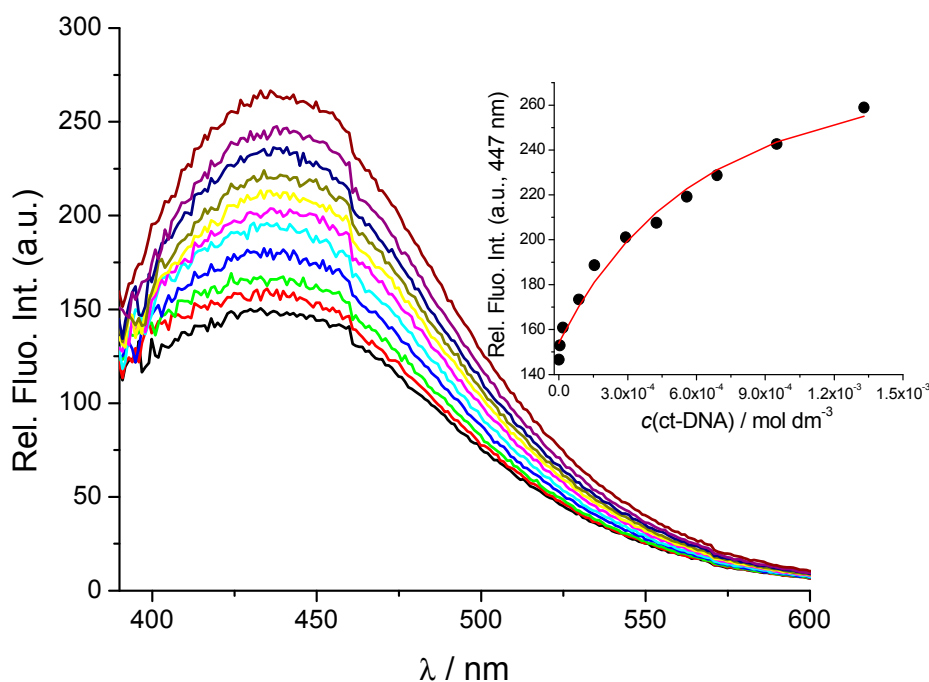


Figure 4. Fluorescence spectra changes of **2** ($c = 2 \times 10^{-6} \text{ mol dm}^{-3}$) upon addition of ct-DNA at pH 7.0, sodium cacodylate buffer, $I = 0.05 \text{ M}$, $\lambda_{\text{exc}} = 355 \text{ nm}$. Inset: emission at 447 nm dependence on $c(\text{DNA})$ fitted to Scatchard equation [40,41].

Table 1. Binding constants ($\log K_s$)^a calculated from the fluorescence titrations ($\lambda_{\text{exc}} = 355$ nm) of **2** and **3** with ds- and ss-polynucleotides at pH = 7.0 and pH = 5.0 (buffer sodium cacodylate, $I = 0.05$ mol dm^{−3}).

Polynucleotide	2		3	
	pH 7	pH 5	pH 7	pH 5
	$\log K_s / {}^c I/I_0$	$\log K_s / {}^c I/I_0$	$\log K_s / {}^c I/I_0$	$\log K_s / {}^c I/I_0$
ct-DNA	3.7/1.8	4.0/1.5	3.7/1.6	4.4/1.5
poly A–poly U	^b	3.5/1.8	^b	^b
poly A	4.1/2.4	^b	^b	^d 4/1.2
poly G	^b	^b	^d 5/0.8	^d 4/0.9
poly C	^b	^b	^b	^b
poly U	^b	^b	^b	^b

^a Processing of titration data by means of Scatchard equation [40,41] gave values of ratio $n[\text{bound } 2, 3]/[\text{polynucleotide}] = 0.1\text{--}0.3$, for easier comparison all $\log K_s$ values were re-calculated for fixed $n = 0.2$. Correlation coefficients were >0.99 for all calculated K_s ; ^b Too small changes for accurate processing by Scatchard eq; ^c I_0 —starting fluorescence intensity of dye; I —fluorescence intensity of dye/polynucleotide complex calculated by Scatchard equation; ^d Small changes allowed only the estimation of $\log K_s$.

In order to gain insight into the changes of polynucleotide properties induced by small molecule binding, we have chosen CD spectroscopy as a highly sensitive method for conformational changes in the secondary structure of polynucleotides [42]. In addition, achiral small molecules can eventually acquire induced CD spectrum (ICD) upon binding to polynucleotides, which could give useful information about modes of interaction [43].

All studied compounds are chiral but intensity of their CD spectra in the 230–500 nm range (Figures S42–S45 of Supplementary Materials) is negligible with respect to CD spectra of polynucleotides, allowing accurate correction of CD titrations. Addition of **2** or **3** to ct-DNA or ds-RNA did not significantly change the CD spectrum of the polynucleotide at both pH 7 and pH 5, thus pointing out that no DNA/RNA conformational change happened upon compound binding. Moreover, no ICD bands >300 nm were observed, thus chromophores of **2**, **3** did not become uniformly oriented along polynucleotide chiral axis [43].

The lack of thermal stabilisation, negligible changes in CD experiments and rather low binding constants ($\log K_s \approx 4$) suggest that the observed fluorescence changes of **2**, **3** are a consequence of random aggregation of compounds along double stranded helix of polynucleotide, based on non-specific electrostatic interactions of their positive charge with DNA/RNA backbone as well as hydrophobic interactions.

Single Stranded (ss)-Polynucleotides

Compounds **2** and **3** contain a fluorescent analogue of cytosine, thus we studied their interactions with ss-polynucleotides. Fluorimetric titrations were performed using sequential excitation properties of **2** and **3**, whereby $\lambda_{\text{exc}} = 299$ nm excites all fluorophores, while excitation at $\lambda_{\text{exc}} = 355$ nm excites only nucleobase-phthalimide. It is important to stress that **1** (reference nucleobase unit) did not change emission upon addition of any ss-polynucleotide at $\lambda_{\text{exc}} = 355$ nm (Figure S40 of Supplementary Materials).

In general, the addition of any ss-polynucleotide to GCP-rigid analogue **2** yielded indiscriminate emission increase upon excitation at $\lambda_{\text{exc}} = 299$ nm. However, excitation of **2** at only nucleobase-fluorophore ($\lambda_{\text{exc}} = 355$ nm) at pH 7 induced significant emission increase of **2** only for poly A (Figure 5a), while intriguingly at pH 5 emission of **2** did not change (Figure S25 of Supplementary Materials), likely due to the formation of poly AH⁺-poly AH⁺ [19,39]. Thus, general fluorescence increase upon $\lambda_{\text{exc}} = 299$ nm can be attributed to the interaction of GCP-unit, which non-selectively binds to all polynucleotides, while highly selective emission upon $\lambda_{\text{exc}} = 355$ nm excitation can be attributed to the particular interaction of nucleobase-unit of **2** with poly A (observed only at pH 7).

For GCP-flexible analogue **3**, the excitation at $\lambda_{\text{exc}} = 299$ nm yielded also fluorescence increase for all ss-RNA. Intriguingly, at pH 7 analogue **3** upon excitation at $\lambda_{\text{exc}} = 355$ nm showed highly selective emission quenching for poly G in respect to other ss-polynucleotides (Figure 5b), but did not show poly A recognition as **2**. Only at acidic conditions (pH 5) at which GCP is protonated (and poly A forms ds-structure poly AH⁺-poly AH⁺) [39], **3** exhibited the poly A-selective response, whereby intriguingly at the same acidic conditions emission of **2** did not change (Figure S35 of Supplementary Materials). Thus, only **3** yielded emission response upon interacting with poly AH⁺-poly AH⁺.

To examine the influence of GCP-unit on interaction with ss-polynucleotides, we tested interactions of compound **12** (analogue of **2** lacking GCP unit) with poly A (Figure S41 of Supplementary Materials), which at both, pH 7 and pH 5 showed non-selective fluorimetric results, similar to the response of free nucleobase **1** (Pyrr-C). These fluorimetric titrations clearly stressed synergistic effect of GCP conjugation with nucleobase for the ss-polynucleotide recognition by **2** and **3**. All fluorimetric titrations with sufficiently large emission change were processed by means of Scatchard eq. to obtain binding constants (Table 1).

In CD experiments addition of **2** or **3** did not change significantly the CD spectra of ss-polynucleotides, at both pH 7 and pH 5, thus indicating that no polynucleotide conformational change happened upon compound binding. Moreover, no ICD bands >300 nm were observed, thus chromophores of **2**, **3** upon binding did not become uniformly oriented along polynucleotide chiral axis [43].

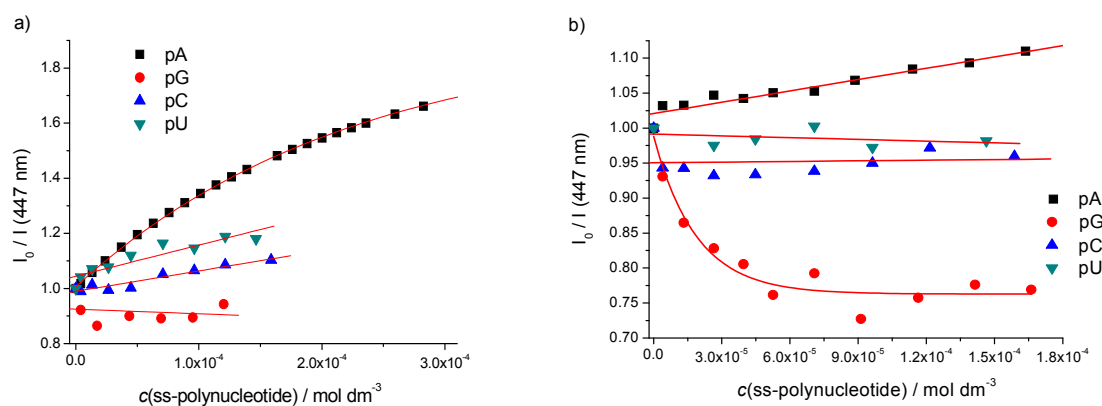


Figure 5. Fluorimetric titrations normalised to starting fluorescence intensity of compound: for **2** (a) and **3** (b), $\lambda_{\text{exc}} = 355$ nm, with all ss-polynucleotides. Done at pH 7.0, sodium cacodylate buffer, $I = 0.05$ M.

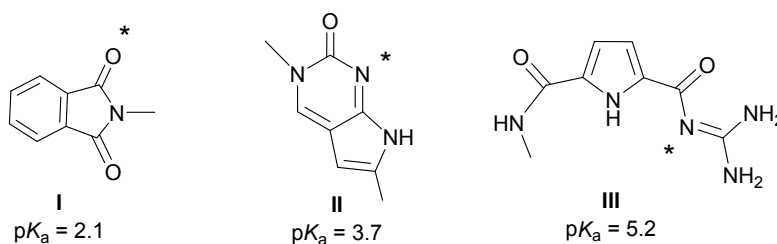
2.4. Molecular Modelling

The ss-RNA do not have well-defined structural features in aqueous solutions and therefore did not allow accurate modelling approach. However, conjugates **2** and **3**, although isosteres, significantly differ in flexibility of GCP-unit attachment, former being significantly more rigid than latter (Figure 1). Such difference in structure could be responsible for dramatic differences in ss-RNA recognition.

In order to examine conformational features of **2** and **3** and inspect whether their intrinsic dynamics in aqueous solution play any role in determining their ability to interact with biological systems studied here, we performed molecular dynamics (MD) simulations of different protonation forms of **2** and **3** placed in the explicit water solvation followed by the quantum-mechanical analysis at the DFT level with implicit solvation.

Initially, we performed the analysis of the acid/base features of **2** and **3** in water to determine their predominant protonation forms under experimental conditions (pH values of 5 and 7). For that purpose, we calculated aqueous phase pK_a values of fragments I–III (Scheme 4) that are building blocks of compounds **2–3**. In doing so, we assumed that the free amino group originating from lysine

will be protonated at both pH values since in lysine its pK_a value is 10.5, which is unlikely to change much or be reduced by 3–4 pK_a units in cases studied here [44].



Scheme 4. Constituting fragments of systems **2** and **3** and the pK_a values calculated at the MP2/6–311++G(2df,2pd)//(SMD)/M06–2X/6–31+G(d) level of theory. The most favourable protonation atom is denoted with an asterisk. * represents the most basic position within a molecule.

The pK_a values calculated for **I** and **II** are slightly lower than those of succinimide ($pK_a = 3.9$) and cytosine ($pK_a = 4.5$) used as references for their evaluation. This implies that even at pH = 5 both of these fragments will predominantly be unprotonated. Interestingly, although guanidine ($pK_a = 13.6$) and its derivatives provide some of the strongest organic bases known today [45–47], combining it to the carbonyl-pyrrole fragment significantly reduces its basicity to $pK_a(\text{III}) = 5.2$, being a consequence of an extended conjugated π -system, which depletes the electron density from the guanidine imino nitrogen leading to much-reduced basicity. This suggests that at pH = 7, the fragment **III** will be neutral, while at pH = 5 it will assume monoprotonated form **III**⁺, thus putting our results in firm agreement with earlier experimental observations [22], and confirming that this fragment is predominantly responsible for the biological activity investigated here.

Having this in mind, we prepared the geometries of monocationic **2**⁺ and **3**⁺ (protonated at the free amino group, corresponding to pH = 7), as well as dicationic **2**²⁺ and **3**²⁺ (also protonated at the guanidine moiety, corresponding to pH = 5), and submitted all four systems to MD simulations. Following the clustering analysis of the obtained trajectories, we optimized a set of representative structures in each case at the (SMD)/M06–2X/6–31G(d) level of theory and extracted total electronic energies in search of most stable, thus most dominant structures in solution (Figure 6).

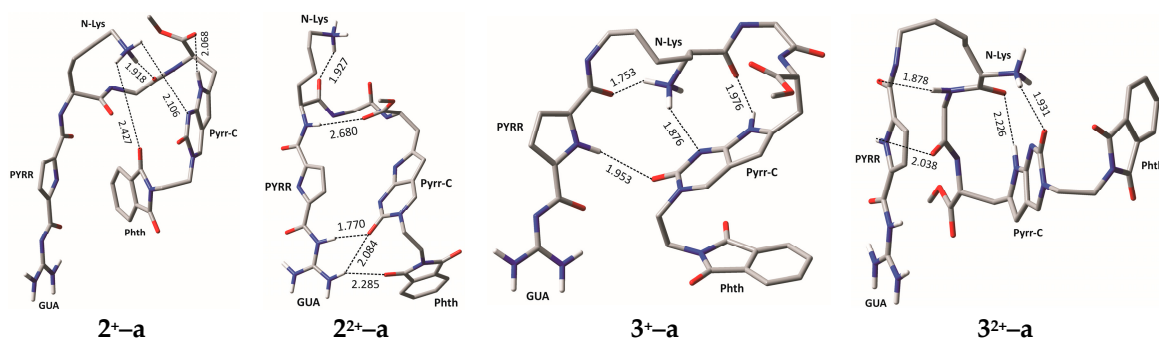


Figure 6. Most stable structures of **2** and **3** in their mono- and di-protonated forms. These are identified after a set of structures, elucidated with the clustering analysis of the corresponding MD trajectories, was optimized at the (SMD)/M06–2X/6–31G(d) level of theory. A larger set of representative structures is placed in the Supplementary Materials (Figures S48–S51). Abbreviations GUA, PYRR, N-Lys, Pyrr-C and Phth denote the guanidine group, pyrrole moiety, free amino group, nucleobase, and phthalimide moiety, respectively. Only selected hydrogen atoms are depicted due to clarity.

It is interesting to observe that, at neutral pH, **2** reveals notable interactions only with ss-poly A, while, surprisingly, at pH = 5, no interactions are detected with either of the ss-polynucleotides. Under

acidic conditions, the most stable structure 2^{2+} -a has the guanidine fragment engaged in three rather strong hydrogen bonds involving one of its amino groups interacting with the carbonyl moiety on the nucleobase (1.77 Å), and the other amino group interacting both with the same carbonyl fragment (2.08 Å) and the carbonyl centre from phthalimide (2.29 Å). This is the only case in all four studied systems where, in the most dominant structure, the guanidine moiety is so strongly involved in the intramolecular stabilization, which is a pattern also observed in several other structures subsequently less stable than 2^{2+} -a (Figure S49 of Supplementary Materials).

This clearly diminishes the ability of the guanidine group in 2^{2+} to interact with other systems as revealed at pH = 5 here. Let us also mention that the most stable 2^{2+} structure having the guanidine group not involved in any intramolecular hydrogen bonding, 2^{2+} -e (Figure S49 of Supplementary Materials), is 2.9 kcal/mol less stable than 2^{2+} -a, which turns out to be too high to be overcome and provides evidence to no activity of **2** towards any of the ss-polynucleotides at pH = 5.

At pH = 7, system **2** loses an extra proton on the guanidine moiety and is present as monocationic 2^+ . In the most stable structure 2^+ -a the guanidine moiety is not involved in any hydrogen bonding, which obviously allows 2^+ to interact with ss-poly A. The intramolecular stabilization in 2^+ -a predominantly occurs through the interaction of the protonated free amino group with (a) the imino nitrogen on the nucleobase (2.11 Å); (b) the carbonyl group on phthalimide (2.43 Å); and (c) the carbonyl moiety in the amide fragment (1.92 Å). Interestingly, only 0.6 kcal mol^{−1} above 2^+ -a, a structure 2^+ -b is also dominated by the intramolecular interactions with the protonated free amino group, but there the guanidine fragment is found in the stabilizing π - π stacking interactions with phthalimide (Figure S48 of Supplementary Materials), which likely affects its interactions with ss-polynucleotides other than ss-poly A. The structure with a non-interacting cytosine nucleobase is 9.6 kcal mol^{−1} less stable than 2^+ -a, which is clearly too high to have any impact on the potential pairing with ss-poly G.

Under acidic conditions (pH = 5), system **3** is the only molecule that exhibits interactions with two polynucleotides, namely, poly A and poly G (Table 1). In its most stable structure 3^{2+} -a the guanidine moiety is not involved in any hydrogen bonding, although the nearby pyrrole fragment and vicinal carbonyl group interact with the neighbouring amide fragments at 2.04 and 1.88 Å, respectively. The structure where the whole GCP fragment is free of any hydrogen bonds is 3^{2+} -b (Figure S51 of Supplementary Materials), which is 2.4 kcal mol^{−1} less stable than 3^{2+} -a. Taken all together, all of this allows 3^{2+} to interact with both poly A and poly G, yet relative changes in increasing or decreasing the fluorescence upon binding are the smallest here (Table 1).

Under physiological conditions (pH = 7, system 3^+), the situation with **3** is similar to that with 2^+ in a way that it shows notable interactions with only one polynucleotide, namely ss-poly G (Table 1). The most stable structure 3^+ -a has the guanidine moiety free of any hydrogen bonding interactions, yet, as in 3^{2+} -a, the interactions of the pyrrole and amide carbonyl of the GCP fragment with the nucleobase carbonyl (1.95 Å) and the protonated free amino group (1.75 Å), respectively, are evident. The structure with the whole GCP fragment deprived of intramolecular interactions is here only 1.9 kcal/mol above the most stable structure, which is likely responsible for slightly higher log K_s = 5 with 3^+ compared to log K_s = 4 with 3^{2+} .

In concluding this section, let us mention that the computational results aided in the interpretation of the relative trends in the observed reactivities of **2** and **3** towards ss-polynucleotides. Since the conformational analysis revealed that in all cases there is not a single structure below 5–6 kcal mol^{−1} having nucleobase fragment not interacting intramolecularly with the rest of the system, this leads us to conclude that in the investigated cases nucleobase moiety is less important for the interaction with ss-polynucleotides. As such, this analysis underlined the important role of the GCP fragment, particularly its guanidine moiety, in successful polynucleotide recognition. It turned out that lowering the pH conditions promotes the interactions with ss-polynucleotides, since protonated guanidine group interacts better, being in agreement with previous experiments [21–25]. Yet, in a more rigid system, this effect is overshadowed by the propensity of 2^{2+} to form strong charge-dipole intramolecular

interactions with protonated guanidine group, therefore hindering its ability to interact with other systems. This is why 2^+ outnumbers 2^{2+} in intermolecular interactions.

3. Materials and Methods

3.1. General Information

Solvents were distilled from appropriate drying agents shortly before use. TLC was carried out on DC-plastikfolien Kieselgel 60 F254 and preparative thick layer (2 mm) chromatography was done on Merck 60 F254 plates (Merck KGaA, Darmstadt, Germany). (Merck, Merck KGaA, Darmstadt, Germany). NMR spectra were recorded on AV600 and AV300 MHz spectrometers (Bruker BioSpin GmbH, Rheinstetten, Germany), operating at 150.92 or 75.47 MHz for ^{13}C and 600.13 or 300.13 MHz for ^1H nuclei using DMSO- d_6 as the internal standard. Mass spectrometry was performed on the Agilent 6410 Triple Quad mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) and high resolution mass spectra (HRMS) were obtained using a Q-Tof2 hybrid quadrupole time-of-flight mass spectrometer (Micromass, Cary, NC, USA). The electronic absorption spectra were obtained on a Cary 100 Bio spectrometer and CD spectra on a JASCO J815 spectrophotometer (JASCO, Oklahoma City, OK, USA), all in quartz cuvettes (1 cm). The spectroscopic studies were performed in aqueous buffer solution (pH 7.0 and pH 5.0, sodium cacodylate buffer, $I = 0.05 \text{ mol dm}^{-3}$). Under the experimental conditions absorbance of **2** and **3** were proportional to its concentration. Polynucleotides were dissolved in sodium cacodylate buffer, $I = 0.05 \text{ mol dm}^{-3}$, pH = 7. Calf thymus (ct-)DNA was additionally sonicated and filtered through a $0.45 \mu\text{m}$ filter [48,49]. Polynucleotide concentration was determined spectroscopically as the concentration of phosphates. Spectroscopic titrations were performed at pH 7.0 and pH 5.0 ($I = 0.05 \text{ mol dm}^{-3}$, sodium cacodylate buffer) by adding portions of polynucleotide solution into the solution of the studied compound for UV/Vis experiments and for CD experiments were done by adding portions of compound stock solution into the solution of polynucleotide. Titration data were processed by Scatchard equation. Values for K_s and n given in Table 1 all have satisfactory correlation coefficients (>0.999). Thermal melting curves for DNA and their complexes with studied compounds were determined as previously described [50] by following the absorption change at 260 nm as a function of temperature. Absorbance of the ligands was subtracted from every curve, and the absorbance scale was normalized. The T_m values are the midpoints of the transition curves, determined from the maximum of the first derivative and checked graphically by the tangent method. ΔT_m values were calculated subtracting T_m of the free nucleic acid from T_m of the complex. Every ΔT_m value here reported was the average of at least two measurements, the error in ΔT_m is $\pm 0.5^\circ\text{C}$.

3.2. Synthesis

***N*⁴-Benzoyl-5-iodocytosine 5:** To a suspension of 5-iodocytosine **4** (2.9 g, 12.24 mmol) and DMAP (196 mg, 1.6 mmol) in anhydrous CH_3CN (26.2 mL) at room temperature, benzoic anhydride (3.05 g, 13.46 mmol) was added. The reaction mixture was heated to reflux for 18 h and then EtOH (50 mL) was added dropwise. After standing at room temperature overnight, the solid was filtered off, washed with EtOH ($3 \times 20 \text{ mL}$) dried in vacuo to afford **5** (3.42 g, 82%) as a white solid: $^1\text{H-NMR}$ (300 MHz, DMSO- d_6) δ /ppm: 12.87 (s, 1H, NH-4), 11.93 (br. s, 1H, NH-1), 8.24 (d, $J = 6.4 \text{ Hz}$, 2H, Bz), 8.17 (s, 1H, H-6), 7.61 (m, 1H, Bz), 7.51 (m, 2H, Bz); $^{13}\text{C-NMR}$ (151 MHz, DMSO- d_6) δ /ppm: 178.1 (C=O, Bz), 158.3 (C-2), 149.9 (C-6), 148.1 (C-4), 136.6 (Bz), 132.6 (Bz), 129.4 (Bz), 128.3 (Bz), 67.6 (C-5). ESI-MS: m/z calcd. for $\text{C}_{11}\text{H}_7\text{IN}_3\text{O}_2 [\text{M} - \text{H}]^-$ 339.96, found 339.80.

***N*⁴-Benzoyl-1-(2-phthalimidoethyl)-5-iodocytosine 6:** *N*⁴-benzoyl-5-iodocytosine **5** (198 mg, 0.58 mmol), potassium carbonate (80.1 mg, 0.58 mmol) and *N*-(2-bromoethyl)phthalimide (294.6 mg, 1.16 mmol) and dry DMF (12 mL) were put in a 20 mL glass vial equipped with a small magnetic stirring bar. Then the mixture was put into the cavity of the MW reactor (Monowave 3000, Anton Paar® GmbH, A-8054 Graz, Austria) and irradiated at 140°C for 35 min. After evaporation of the solvent, CH_2Cl_2

and MeOH were added and the resulting solid was filtered off. The filtrate was evaporated and the residue was purified by preparative chromatography (hexane/ethyl acetate 1:1) to give 152 mg (51%) of yellow powder **6**: $^1\text{H-NMR}$ (300 MHz, $\text{DMSO-}d_6$) δ /ppm: 12.87 (br. s, 1H, NH-4), 8.41 (s, 1H, H-6), 8.14 (br. s, 2H, Bz), 7.90–7.82 (m, 4H, Pht), 7.58–7.46 (m, 3H, Bz), 4.05–3.95 (br. s, 2H, CH_2Pht), 3.94–3.86 (br. s, 2H, CH_2Pht); $^{13}\text{C-NMR}$ (75 MHz, $\text{DMSO-}d_6$) δ /ppm: 173.8 (C=O, Bz), 167.7 (C=O, Pht), 163.9 (C-2), 155.9 (C-4), 151.8 (C-6), 137.9 (Bz), 134.4 (Pht), 134.3 (Bz), 131.6 (Pht), 129.2 (Bz), 128.0 (Bz), 123.1 (Pht), 59.9 (C-5), 47.5 (CH_2Pht), 36.6 (CH_2Pht). ESI-MS: m/z calcd. for $\text{C}_{21}\text{H}_{14}\text{IN}_4\text{O}_4$ $[\text{M} - \text{H}]^-$ 513.01, found 512.60.

Pyrr-C Amino acid analogue 8: To a solution of N^4 -benzoyl-1-(2-phthalimidoethyl)-5-iodocytosine **6** (40 mg, 0.078 mmol) in anhydrous DMF (546 μL), CuI (3 mg, 0.016 mmol), Et_3N (22 μL , 0.156 mmol), protected propargylglycine **7** (53.2 mg, 0.234 mmol) and $\text{Pd}(\text{PPh}_3)_4$ (9 mg, 0.008 mmol) were sequentially added. The reaction mixture was stirred in the dark at 50 $^\circ\text{C}$ for 24 h after which the solvent was removed under vacuum. The residue was purified by preparative chromatography ($\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ 9:1) to give the product **8** (25 mg, 64%) as a yellow powder: $^1\text{H-NMR}$ (300 MHz, $\text{DMSO-}d_6$) δ /ppm: 10.99 (s, 1H, NH-7), 8.27 (s, 1H, H-4), 7.81 (s, 4H, Pht), 7.32 (d, $J = 8.0$ Hz, 1H, NH), 5.85 (s, 1H, H-5), 4.38–4.30 (m, 1H, CHCH_2), 4.17–4.11 (m, 2H, CH_2Pht), 3.95–3.92 (m, 2H, CH_2Pht), 3.60 (s, 3H, OCH_3), 2.97–2.73 (m, 2H, CH_2CH), 1.38–1.23 (m, 9H, $\text{C}(\text{CH}_3)_3$); $^{13}\text{C-NMR}$ (75 MHz, $\text{DMSO-}d_6$) δ /ppm: 172.2 (C=O), 167.7 (C=O, Pht), 159.5 (C-2), 155.3 (C=O), 154.7 (C-6), 140.3 (C-4), 137.2 (C-7a), 134.3 (Pht), 131.6 (Pht), 123.1 (Pht), 108.6 (C-4a), 98.1 (C-5), 78.4 ($\text{C}(\text{CH}_3)_3$), 52.5 (CHCH_2), 51.9 (OCH_3), 49.3 (CH_2Pht), 37.0 (CH_2Pht), 29.4 (CH_2CH), 28.1 ($\text{C}(\text{CH}_3)_3$). ESI-MS: m/z calcd. for $\text{C}_{25}\text{H}_{28}\text{N}_5\text{O}_7$ $[\text{M} + \text{H}]^+$ 510.20, found 510.30.

Pyrr-C Amino acid analogue 1: Compound **8** (160 mg, 0.314 mmol) was dissolved in a 1:1 mixture of TFA/ CH_2Cl_2 (4 mL) and stirred at room temperature for 72 h. After evaporation of volatiles, the residue was purified by preparative chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$ 9:1). Product **1** (144 mg, 88%) was obtained as a yellow foam: $^1\text{H-NMR}$ (300 MHz, $\text{DMSO-}d_6$) δ /ppm: 10.94 (br. s, 1H, NH-7), 8.24 (s, 1H, H-4), 7.82 (s, 4H, Pht), 7.70–7.31 (m, 2H, NH_2), 5.84 (s, 1H, H-5), 4.14–4.12 (m, 2H, CH_2Pht), 3.95–3.93 (m, 2H, CH_2Pht), 3.74–3.65 (m, 1H, CHCH_2), 3.58 (s, 3H, OCH_3), 2.89–2.69 (m, 2H, CH_2CH); $^{13}\text{C-NMR}$ (75 MHz, $\text{DMSO-}d_6$) δ /ppm: 175.1 (C=O), 167.7 (C=O), 159.5 (C-2), 154.8 (C-6), 140.0 (C-4), 138.2 (C-7a), 134.3 (Pht), 131.6 (Pht), 123.1 (Pht), 108.7 (C-4a), 97.7 (C-5), 53.5 (CHCH_2), 51.5 (OCH_3), 49.3 (CH_2Pht), 37.0 (CH_2Pht), 33.2 (CH_2CH). HRMS (MALDI-TOF/TOF): m/z calcd. for $\text{C}_{20}\text{H}_{20}\text{N}_5\text{O}_5$ $[\text{M}]^+$ 410.1464, found 410.1458.

Z-D-Lys(Boc)-Gly-OMe 9: Z-D-Lys(Boc)OH (300 mg, 0.789 mmol) and H-Gly-OMe hydrochloride (98.6 mg, 0.789 mmol) were dissolved in dry CH_3CN (10 mL) under argon and HOBt (105.4 mg, 0.789 mmol), HBTU (295.8 mg, 0.789 mmol) and dry Et_3N (435 μL , 3.16 mmol) were added. The reaction mixture was stirred at room temperature overnight. Product **9** (348 mg, 98%) was isolated by preparative chromatography ($\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ 9:1) as a white powder: $^1\text{H-NMR}$ (300 MHz, $\text{DMSO-}d_6$) δ /ppm: 8.31 (t, $J = 5.8$ Hz, 1H, NH), 7.40–7.29 (m, 6H, Ph, NH), 6.75 (t, $J = 5.1$ Hz, 1H, NH), 5.02 (s, 2H, OCH_2Ph), 4.02–3.95 (m, 1H, CHCH_2), 3.92–3.75 (m, 2H, COCH_2NH), 3.62 (s, 3H, OCH_3), 2.91–2.84 (m, 2H, CH_2NH), 1.68–1.49 (m, 2H, CH_2CH), 1.46–1.23 (m, 13H, $2 \times \text{CH}_2$, $\text{C}(\text{CH}_3)_3$); $^{13}\text{C-NMR}$ (75 MHz, $\text{DMSO-}d_6$) δ /ppm: 172.6 (C=O), 170.2 (C=O), 156.0 (C=O, Cbz), 155.6 (C=O, Boc), 137.0 (Ph), 128.3 (Ph), 127.8 (Ph), 127.7 (Ph), 77.4 ($\text{C}(\text{CH}_3)_3$), 65.4 (OCH_2Ph), 54.5 (CHCH_2), 51.7 (OCH_3), 40.5 (COCH_2NH), 39.8 (CH_2NH), 31.6 (CH_2CH), 29.2 (CH_2), 28.3 ($\text{C}(\text{CH}_3)_3$), 22.7 (CH_2). ESI-MS: m/z calcd. for $\text{C}_{22}\text{H}_{33}\text{N}_3\text{NaO}_7$ $[\text{M} + \text{Na}]^+$ 474.22, found 474.22.

Z-D-Lys(Boc)-Gly-OH 10: Compound **9** (170 mg, 0.377 mmol) was dissolved in a 3:1 mixture of dioxane/water (4 mL), 2M NaOH (470 μL , 0.942 mmol) was added and reaction mixture was stirred at room temperature for 24 h. The solvent was evaporated, the residue was dissolved in water (~20 mL) and acidified with 1 M HCl to pH 5. The product was extracted with EtOAc (5×20 mL) dried over Na_2SO_4 yielding **10** (90 mg, 55%) as a colourless soil: $^1\text{H-NMR}$ (600 MHz, $\text{DMSO-}d_6$) δ /ppm: 7.55

(d, $J = 8.1$ Hz, 1H, NH), 7.37–7.23 (m, 6H, Ph, NH), 6.75–6.73 (m, 1H, NH), 5.03 (q, $J = 12.6$ Hz, 2H, OCH_2Ph), 3.90–3.86 (m, 1H, CHCH_2), 3.25–3.21 (m, 2H, COCH_2NH), 2.89–2.83 (m, 2H, CH_2NH), 1.64–1.24 (m, 15H, $3 \times \text{CH}_2$, $\text{C}(\text{CH}_3)_3$); ^{13}C -NMR (75 MHz, $\text{DMSO}-d_6$) δ /ppm: 170.6 (C=O), 170.0 (C=O), 156.1 (C=O, Cbz), 155.6 (C=O, Boc), 137.1 (Ph), 128.4 (Ph), 127.7 (Ph), 127.6 (Ph), 77.3 ($\text{C}(\text{CH}_3)_3$), 65.3 (OCH_2Ph), 55.1 (CHCH_2), 44.1 (COCH_2NH), 39.9 (CH_2NH), 31.4 (CH_2CH), 29.1 (CH_2), 28.3 ($\text{C}(\text{CH}_3)_3$), 22.8 (CH_2). ESI-MS: m/z calcd. for $\text{C}_{21}\text{H}_{30}\text{N}_3\text{O}_7$ $[\text{M} - \text{H}]^-$ 436.21, found 436.0.

Pyrr-C Tripeptide 11: Pyrr-C amino acid analogue **1** (46 mg, 0.088 mmol) and dipeptide **10** (38.4 mg, 0.088 mmol) were dissolved in dry CH_3CN (5 mL) under argon and HOBt (11.9 mg, 0.088 mmol), HBTU (33.3 mg, 0.088 mmol) and dry Et_3N (49 μL , 0.352 mmol) were added. Reaction was stirred at room temperature overnight. Product **11** (36 mg, 49%) was isolated by preparative chromatography ($\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ 9:1) as a yellow powder: ^1H -NMR (300 MHz, $\text{DMSO}-d_6$) δ /ppm: 11.03 (br. s, 1H, NH-7), 8.34–8.32 (m, 1H, NH), 8.26 (s, 1H, H-4), 8.17–8.12 (m, 1H, NH), 7.81 (s, 4H, Pht), 7.48–7.42 (m, 1H, NH), 7.35–7.26 (m, 5H, Ph), 6.75 (t, $J = 5.6$ Hz, 1H, NH), 5.87 (s, 1H, H-5), 5.05–4.95 (m, 2H, OCH_2Ph), 4.62 (br. s, 1H, CHCH_2), 4.20–4.08 (m, 2H, CH_2 Pht), 4.00–3.89 (m, 3H, CHCH_2 , CH_2Pht), 3.74–3.65 (m, 2H, COCH_2NH), 3.58 (s, 3H, OCH_3) 2.97–2.85 (m, 4H, CH_2CH , CH_2NH), 1.62–1.49 (m, 2H, CH_2CH), 1.44–1.16 (m, 15H, $3 \times \text{CH}_2$, $\text{C}(\text{CH}_3)_3$); ^{13}C -NMR (151 MHz, $\text{DMSO}-d_6$) δ /ppm: 172.3 (C=O), 171.4 (C=O), 168.8 (C=O), 167.7 (C=O), 159.5 (C-2), 156.1 (C=O, Cbz), 155.6 (C=O, Boc), 154.8 (C-6), 140.3 (C-4), 139.0 (C-7a), 136.8 (Ph), 134.3 (Pht), 131.6 (Pht), 128.3 (Ph), 127.8 (Ph), 127.7 (Ph), 123.1 (Pht), 108.6 (C-4a), 98.1 (C-5), 77.3 ($\text{C}(\text{CH}_3)_3$), 65.4 (OCH_2Ph), 54.7 (CHCH_2), 52.0 (OCH_3), 51.2 (CHCH_2), 49.3 (CH_2Pht), 41.6 (COCH_2NH), 39.8 (CH_2NH), 37.0 (CH_2Pht), 31.4 (CH_2CH), 29.6 (CH_2), 29.2 (CH_2), 28.3 ($\text{C}(\text{CH}_3)_3$), 22.8 (CH_2). ESI-MS: m/z calcd. for $\text{C}_{41}\text{H}_{48}\text{N}_8\text{NaO}_{11}$ $[\text{M} + \text{Na}]^+$ 851.33, found 851.50.

Pyrr-C Tripeptide 12: Compound **11** (160 mg, 0.193 mmol) was dissolved in 1:1 mixture of TFA/ CH_2Cl_2 (4 mL) and stirred at room temperature for 20 h. After removal of remaining TFA under reduced pressure, product **12** (162.7 mg, 100%) was obtained as a yellow foam: ^1H -NMR (300 MHz, $\text{DMSO}-d_6$) δ /ppm: 11.19 (br. s, 1H, NH-7), 8.38–8.35 (m, 2H, NH, H-4), 8.19–8.14 (m, 1H, NH), 7.81 (s, 4H, Pht), 7.63 (br. s, 3H, NH_3), 7.50–7.46 (m, 1H, NH), 7.36–7.31 (m, 5H, Ph), 5.92 (s, 1H, H-5), 5.04–4.97 (m, 2H, OCH_2Ph), 4.64–4.61 (m, 1H, CHCH_2), 4.20–3.59 (m, 10H, $2 \times \text{CH}_2\text{Pht}$, CHCH_2 , COCH_2NH , $\text{OCH}_3 + \text{H}_2\text{O}$), 3.02–2.86 (m, 2H, CH_2CH), 2.74 (br. s, 2H, CH_2), 1.62–1.17 (m, 6H, $3 \times \text{CH}_2$); ^{13}C -NMR (151 MHz, $\text{DMSO}-d_6$) δ /ppm: 172.5 (C=O), 172.4 (C=O), 172.4 (C=O), 168.0 (C=O), 159.8 (C-2), 158.5, 158.4, 158.3, 157.8 ($4 \times \text{CF}_3\text{C=O}$), 156.4 (C=O, Cbz), 154.5 (C-6), 140.8 (C-4), 139.9 (C-7a), 137.3 (Ph), 134.6 (Pht), 131.7 (Pht), 128.6 (Ph), 127.9 (Ph), 123.3 (Pht), 119.7 ($\text{CF}_3\text{C=O}$), 109.2 (C-4a), 98.9 (C-5), 66.0 (OCH_2Ph), 52.4 (CHCH_2), 51.53 (OCH_3), 51.51 (CHCH_2), 49.7 (CH_2Pht), 40.2 (COCH_2NH), 38.9 (CH_2NH), 37.2 (CH_2Pht), 31.2 (CH_2CH), 29.9 (CH_2), 26.7 (CH_2), 22.4 (CH_2). ESI-MS: m/z calcd. for $\text{C}_{36}\text{H}_{41}\text{N}_8\text{O}_9$ $[\text{M}]^+$ 729.3, found 729.3.

Pyrr-C Tripeptide GCP conjugate 14: Compound **12** (30 mg, 0.035 mmol) and Boc-GCP-OH **13** (14.1 mg, 0.035 mmol) were dissolved in dry CH_3CN (3 mL) under argon and HOBt (4.7 mg, 0.035 mmol), HBTU (13.3 mg, 0.035 mmol) and dry Et_3N (19 μL , 0.14 mmol) were added. Reaction was stirred at room temperature overnight. Product **14** (20 mg, 57%) was isolated by preparative chromatography ($\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ 9:1) as a yellow powder: ^1H -NMR (300 MHz, $\text{DMSO}-d_6$) δ /ppm: 11.03 (m, 2H, NH-7, NH), 9.27 (br. s, 2H, $2 \times \text{NH}$), 8.59–8.42 (m, 1H, NH), 8.40–8.29 (m, 1H, NH), 8.27 (s, 1H, H-4), 8.18 (m, 1H, NH), 7.80 (s, 4H, Pht), 7.50 (m, 1H, NH), 7.34–7.20 (m, 6H, Ph, NH), 6.79–6.67 (m, 2H, CH-Pyrr), 5.87 (s, 1H, H-5), 5.03–4.96 (m, 2H, OCH_2Ph), 4.63–4.59 (m, 1H, CHCH_2), 4.18–4.09 (m, 2H, CH_2Pht), 4.01–3.92 (m, 3H, CHCH_2 , CH_2Pht), 3.77–3.67 (m, 2H, COCH_2NH), 3.58 (s, 3H, OCH_3), 3.19–3.16 (m, 2H, CH_2NH), 3.00–2.86 (m, 1H, CH_2CH), 1.63–1.17 (m, 15H, $3 \times \text{CH}_2$, $\text{C}(\text{CH}_3)_3$); ^{13}C -NMR (151 MHz, $\text{DMSO}-d_6$) δ /ppm: 172.9 (C=O), 172.1 (C=O), 171.8 (C=O), 170.1 (C=O), 169.4 (C=O), 168.2 (C=O), 160.2, 159.7 (C-2), 158.8, 156.7 (C=O, Cbz), 155.3 (C=O, Boc), 154.4 (C=O, C-6), 140.8 (C-4), 137.3 (C-7a), 134.8 (Pht), 131.8 (Pht), 128.7 (Ph), 127.8 (Ph), 123.5 (Pht), 121.0 (Pyrr), 112.2 (CH-Pyrr), 109.3 (C-4a), 98.7 (C-5), 78.1 ($\text{C}(\text{CH}_3)_3$), 65.9 (OCH_2Ph), 55.1 (CHCH_2), 52.5 (OCH_3), 51.6

(CHCH₂), 49.7 (CH₂Pht), 42.0 (COCH₂NH), 38.9 (CH₂NH), 37.4 (CH₂Pht), 31.6 (CH₂CH), 29.9 (CH₂), 29.0 (CH₂), 28.1 (C(CH₃)₃), 23.2 (CH₂). ESI-MS: *m/z* calcd. for C₄₈H₅₃N₁₂O₁₃ [M – H][–] 1005.39, found 1004.9.

Lys(Boc)-Gly-OMe 16: Z-D-Lys(Boc)-Gly-OMe **9** (430 mg, 0.95 mmol) was dissolved in methanol (50 mL) and 10% Pd/C (43 mg) was added. The reaction mixture was stirred under a hydrogen atmosphere for 24 h at room temperature, after which time the catalyst was filtered off through a Celite pad and washed several times with methanol. The solvent was evaporated under reduced pressure to provide the product **16** (302 mg, 100%) as a colourless oil: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 8.50–8.46 (m, 1H, NH), 7.64–7.56 (m, 1H, NH), 6.79–6.73 (m, 2H, NH₂), 4.09 (br. s, 1H, CHCH₂), 3.96–3.82 (m, 2H, COCH₂NH), 3.63 (s, 3H, OCH₃), 2.96–2.82 (m, CH₂NH), 1.69–1.56 (m, 2H, CH₂CH), 1.42–1.33 (m, 13H, 2 × CH₂, C(CH₃)₃); ¹³C-NMR (75 MHz, DMSO-*d*₆) δ/ppm: 170.2 (C=O), 168.0 (C=O), 166.1 (C=O), 155.6 (C=O), 77.4 (C(CH₃)₃), 54.1 (CHCH₂), 51.7 (OCH₃), 44.3 (COCH₂NH), 40.5 (CH₂NH), 33.3 (CH₂CH), 29.3 (CH₂), 28.3 (C(CH₃)₃), 22.0 (CH₂). ESI-MS: *m/z* calcd. for C₁₄H₂₈N₃O₅ [M + H]⁺ 318.2, found 318.3.

GCP Dipeptide 17: Dipeptide **16** (61 mg, 0.192 mmol) and Boc-GCP-OH **13** (76.2 mg, 0.192 mmol) were dissolved in dry CH₃CN (3 mL) under argon and HOBt (26 mg, 0.192 mmol), HBTU (73 mg, 0.192 mmol) and dry Et₃N (107 μL, 0.768 mmol) were added. Reaction was stirred at room temperature overnight. Product **17** (67 mg, 58%) was isolated by preparative chromatography (CH₂Cl₂/CH₃OH 9:1) as a yellow powder: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 11.62 (br. s, 1H, NH), 10.88 (br. s, 1H, NH), 9.33 (s, 1H, NH), 8.57 (s, 1H, NH), 8.46–8.41 (m, 2H, 2 × NH), 6.83–6.80 (m, 2H, CH-Pyrr), 6.75 (t, *J* = 5.0 Hz, 1H, NH), 4.47–4.40 (m, 1H, CHCH₂), 3.85–3.80 (m, 2H, COCH₂NH), 3.62 (s, 3H, OCH₃), 2.90–2.88 (m, 2H, CH₂NH), 1.75–1.64 (m, 2H, CH₂CH), 1.61–1.23 (m, 22H, 2 × CH₂, 2 × C(CH₃)₃); ¹³C-NMR (75 MHz, DMSO-*d*₆) δ/ppm: 172.4 (C=O), 170.2 (C=O), 159.4, 158.4, 155.5 (C=O, Boc), 113.6 (CH-Pyrr), 112 (CH-Pyrr), 77.3 (C(CH₃)₃), 52.5 (CHCH₂), 51.6 (OCH₃), 40.5 (COCH₂NH), 39.8 (CH₂NH), 31.6 (CH₂CH), 29.2 (CH₂), 28.2 (C(CH₃)₃), 27.8 (C(CH₃)₃), 22.9 (CH₂). ESI-MS: *m/z* calcd. for C₂₆H₄₂N₇O₉ [M + H]⁺ 596.3, found 596.3.

GCP Dipeptide 18: To a solution of dipeptide **17** (90 mg, 0.147 mmol) in a 4:1 mixture of THF/water (25 mL), lithium hydroxide (37.8 mg, 0.882 mmol) was added and reaction mixture was stirred at room temperature overnight. The solvent was evaporated and a concentrated citric acid solution was added to the water phase (pH ~5). White precipitate which formed was filtered and washed several times with water yielding product **18** (64 mg, 75%) as a white soil: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 11.64 (br. s, 3H, 2 × NH, COOH), 9.32 (s, 1H, NH), 8.57 (s, 1H, NH), 8.45 (d, *J* = 8.1 Hz, 1H, NH), 8.24 (t, *J* = 5.7 Hz, 1H, NH), 6.83 (s, 2H, CH-Pyrr), 6.74 (t, *J* = 5.5 Hz, 1H, NH), 4.45–4.41 (m, 1H, CHCH₂), 3.80–3.66 (m, 2H, COCH₂NH), 2.89–2.87 (m, 2H, CH₂NH), 1.83–1.63 (m, 2H, CH₂CH), 1.60–1.24 (m, 22H, 2 × CH₂, 2 × C(CH₃)₃); ¹³C-NMR (75 MHz, DMSO-*d*₆) δ/ppm: 172.1 (C=O), 171.1 (C=O), 159.4, 158.4, 155.5 (C=O, Boc), 129.4 (Pyrr), 113.7 (CH-Pyrr), 112.9 (CH-Pyrr), 77.3 (C(CH₃)₃), 52.5 (CHCH₂), 40.4 (COCH₂NH), 39.8 (CH₂NH), 31.6 (CH₂CH), 29.2 (CH₂), 28.2 (C(CH₃)₃), 27.8 (C(CH₃)₃), 22.9 (CH₂). ESI-MS: *m/z* calcd. for C₂₅H₄₀N₇O₉ [M + H]⁺ 582.29, found 582.4.

Pyrr-C Tripeptide GCP conjugate 19: Compound **1** (27 mg, 0.052 mmol) and dipeptide **18** (30 mg, 0.052 mmol) were dissolved in dry CH₃CN (5 mL) under argon and HOBt (7 mg, 0.052 mmol), HBTU (19.6 mg, 0.052 mmol) and dry Et₃N (29 μL, 0.208 mmol) were added. Reaction was stirred at room temperature overnight. Product **19** (25 mg, 50%) was isolated by preparative chromatography (CH₂Cl₂/CH₃OH 9:1) as a yellow powder: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 11.01 (br. s, 2H, NH, NH-7), 9.29 (br. s, 2H, 2 × NH), 8.61–8.23 (m, 5H, 4 × NH, H-4), 7.81 (s, 4H, Pht), 6.80–6.73 (m, 3H, NH, 2 × CH-Pyrr), 5.87 (d, *J* = 12.2 Hz, 1H, H-5), 4.64–4.57 (m, 1H, CHCH₂), 4.33 (br. s, 1H, CHCH₂), 4.20–4.07 (m, 2H, CH₂Pht), 3.97–3.88 (m, 2H, CH₂Pht), 3.74–3.65 (m, 2H, COCH₂NH), 3.57 (m, 3H, OCH₃), 3.03–2.87 (m, 4H, CH₂CH, CH₂NH), 1.80–1.56 (m, 2H, CH₂CH), 1.49–1.23 (m, 22H, 2 × CH₂, 2 × C(CH₃)₃); ¹³C-NMR (75 MHz, DMSO-*d*₆) δ/ppm: 172.1 (C=O), 171.4 (C=O), 169.9

(C=O), 168.9 (C=O), 167.7 (C=O), 167.0, 159.5 (C-2), 158.5, 155.5 (C=O, Boc), 154.7 (C-6), 140.3 (C-4), 136.8 (C-7a), 134.3 (Pht), 131.6 (Pht), 123.1 (Pht), 113.1 (CH-Pyrr), 111.9 (CH-Pyrr), 108.6 (C-4a), 98.1 (C-5), 77.3 (C(CH₃)₃), 68.3 (OCH₂Ph), 53.1 (CHCH₂), 52.0 (OCH₃), 51.1 (CHCH₂), 49.3 (CH₂Pht), 41.6 (COCH₂NH), 39.7 (CH₂NH), 37.0 (CH₂Pht), 31.3 (CH₂CH), 29.6 (CH₂), 29.2 (CH₂), 28.2 (C(CH₃)₃), 27.8 (C(CH₃)₃), 22.9 (CH₂). ESI-MS: *m/z* calcd. for C₄₅H₅₅N₁₂O₁₃ [M – H][–] 971.4, found 971.0.

Conjugate Pyrr-C Tripeptide 2: Compound **19** (14 mg, 0.014 mmol) was dissolved in 1:1 mixture of TFA/CH₂Cl₂ (4 mL) and stirred at room temperature for 20 h. After removal of remaining TFA under reduced pressure, product **2** (15 mg, 100%) was obtained as a yellow foam: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 12.57–12.28 (m, 1H, NH), 11.37 (br. s, 1H, NH), 10.98–10.92 (m, 1H, NH-7), 8.59–8.26 (m, 8H, 4 × NH, NH₃, H-4), 7.81–7.70 (m, 7H, Pht, NH₃), 7.12 (s, 1H, CH-Pyrr), 6.89 (s, 1H, CH-Pyrr), 5.91–5.88 (m, 1H, H-5), 4.63–4.61 (m, 1H, CHCH₂), 4.43 (s, 1H, CHCH₂), 4.18–4.13 (m, 2H, CH₂Pht), 3.94 (m, 2H, CH₂Pht), 3.76–3.67 (m, 2H, COCH₂NH), 3.59 (s, 3H, OCH₃), 3.01–2.77 (m, 4H, CH₂CH, CH₂), 1.76–1.23 (m, 6H, 3 × CH₂); ¹³C-NMR (151 MHz, DMSO-*d*₆) δ/ppm: 171.8 (C=O), 171.4 (C=O), 168.8 (C=O), 167.7 (C=O), 159.4 (C-2), 159.1, 158.4, 158.2, 158.0, 157.8 (4 × CF₃C=O), 154.8 (C-6), 140.4 (C-4), 136.9 (C-7a), 134.3 (Pht), 131.6 (Pht), 123.1 (Pht), 118.2 (CF₃C=O), 114.9 (CH-Pyrr), 113.6 (CH-Pyrr), 108.7 (C-4a), 98.3 (C-5), 68.3 (OCH₂Ph), 52.7 (CHCH₂), 52.0 (CHCH₂), 51.3 (OCH₃), 49.3 (CH₂Pht), 41.5 (COCH₂NH), 38.7 (CH₂NH), 37.0 (CH₂Pht), 31.2 (CH₂CH), 29.6 (CH₂), 26.6 (CH₂), 22.4 (CH₂). HRMS (MALDI-TOF/TOF): *m/z* calcd. for C₃₅H₄₀N₁₂O₉ [M + H]⁺ 773.3119; found 773.3137.

Boc-D-Lys(Z)-Gly-OMe 20: Boc-D-Lys(Z)-OH (400 mg, 1.051 mmol) and H-Gly-OMe hydrochloride (131.5 mg, 1.051 mmol) were dissolved in dry CH₃CN (10 mL) under argon and HOBt (141.9 mg, 1.051 mmol), HBTU (398.2 mg, 1.051 mmol) and dry Et₃N (585 μL, 4.2 mmol) were added. Reaction was stirred at room temperature overnight. Product **20** (455 mg, 96%) was isolated by preparative chromatography (CH₂Cl₂/CH₃OH 9:1) as a colorless oil: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 8.21 (t, *J* = 5.8 Hz, 1H, NH), 7.39–7.32 (m, 5H, Ph), 7.22 (t, *J* = 5.3 Hz, 1H, NH), 6.83 (d, *J* = 8.1 Hz, 1H, NH), 5.00 (s, 2H, OCH₂Ph), 3.92–3.74 (m, 3H, CHCH₂, COCH₂NH), 3.61 (s, 3H, OCH₃), 3.00–2.96 (m, 2H, CH₂NH), 1.47–1.24 (m, 15H, 3 × CH₂, C(CH₃)₃); ¹³C-NMR (75 MHz, DMSO-*d*₆) δ/ppm: 172.8 (C=O), 170.3 (C=O), 156.1 (C=O, Cbz), 155.3 (C=O, Boc), 137.3 (Ph), 128.4 (Ph), 127.7 (Ph), 78.0 (C(CH₃)₃), 65.1 (OCH₂Ph), 54.1 (CHCH₂), 51.6 (OCH₃), 40.5 (COCH₂NH), 40.1 (CH₂NH), 31.6 (CH₂CH), 29.1 (CH₂), 28.2 (C(CH₃)₃), 22.7 (CH₂). ESI-MS: *m/z* calcd. for C₂₂H₃₃N₃NaO₇ [M + Na]⁺ 474.22, found 474.22.

Boc-D-Lys-Gly-OMe 21: Cbz-protected compound **20** (380 mg, 0.842 mmol) was dissolved in methanol (50 mL) and 10% Pd/C (38 mg) was added. The reaction mixture was stirred under the hydrogen for 16 h at room temperature, after which time the catalyst was filtered off through a Celite pad and washed several times with methanol. The solvent was evaporated under reduced pressure to provide the product **21** (267 mg, 100%) as a colourless oil: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 8.29 (t, *J* = 5.8 Hz, 1H, NH), 6.89 (d, *J* = 8.1 Hz, 1H, NH), 4.50 (br. s, 2H, NH₂), 3.96–3.74 (m, 3H, CHCH₂, COCH₂NH), 3.62 (s, 3H, OCH₃), 2.60 (t, *J* = 6.9 Hz, 2H, CH₂NH₂), 1.66–1.16 (m, 15H, 3 × CH₂, C(CH₃)₃); ¹³C-NMR (75 MHz, DMSO-*d*₆) δ/ppm: 172.5 (C=O), 167.0 (C=O), 155.0 (C=O, Boc), 77.7 (C(CH₃)₃), 53.7 (CHCH₂), 51.4 (OCH₃), 40.5 (COCH₂NH), 39.8 (CH₂NH), 31.2 (CH₂CH), 29.3 (CH₂), 27.9 (C(CH₃)₃), 22.2 (CH₂). ESI-MS: *m/z* calcd. for C₁₄H₂₈N₃O₅ [M + H]⁺ 318.2, found 318.2.

GCP Dipeptide 22: Compound **21** (103 mg, 0.325 mmol) and Boc-GCP-OH **13** (128.9 mg, 0.325 mmol) were dissolved in dry CH₃CN (8 mL) under argon and HOBt (43.9 mg, 0.325 mmol), HBTU (123.2 mg, 0.325 mmol) and dry Et₃N (181 μL, 1.3 mmol) were added. Reaction was stirred at room temperature overnight. Product **22** (78 mg, 40%) was isolated by preparative chromatography (eluent: CH₂Cl₂/CH₃OH 9:1) as a slightly yellow oil: ¹H-NMR (300 MHz, DMSO-*d*₆) δ/ppm: 11.30 (br. s, 2H, 2 × NH), 9.27 (s, 1H, NH), 8.53–8.22 (m, 3H, 3 × NH), 6.86 (d, *J* = 8.2 Hz, 1H, NH), 6.75 (d, *J* = 3.6 Hz, 1H, CH-Pyrr), 6.66 (d, *J* = 3.6 Hz, 1H, CH-Pyrr), 3.96–3.74 (m, 3H, CHCH₂, COCH₂NH), 3.61 (s, 3H, OCH₃), 3.22–3.16 (m, 2H, CH₂NH), 1.50–1.29 (m, 24H, 3 × CH₂, 2 × C(CH₃)₃); ¹³C-NMR (75 MHz, DMSO-*d*₆) δ/ppm: 172.8 (C=O), 170.2 (C=O), 158.5, 155.2 (C=O, Boc), 113.6 (CH-Pyrr), 111.9 (CH-Pyrr),

77.9 ($\underline{\text{C}}(\text{CH}_3)_3$), 54.0 ($\underline{\text{CHCH}_2}$), 51.6 (OCH_3), 40.5 (COCH_2NH), 38.3 ($\underline{\text{CH}_2\text{NH}}$), 31.6 ($\underline{\text{CH}_2\text{CH}}$), 29.0 (CH_2), 28.2 ($\text{C}(\underline{\text{CH}_3})_3$), 27.8 ($\text{C}(\underline{\text{CH}_3})_3$), 22.9 (CH_2). ESI-MS: m/z calcd. for $\text{C}_{26}\text{H}_{42}\text{N}_7\text{O}_9$ $[\text{M} + \text{H}]^+$ 596.3, found 596.3.

GCP Dipeptide 23: To a solution of **22** (69 mg, 0.116 mmol) in a 4:1 mixture of THF/water (25 mL), lithium hydroxide (29.2 mg, 0.696 mmol) was added and reaction mixture was stirred at room temperature overnight. The solvent was evaporated and a concentrated citric acid solution was added to the water phase (pH ~5). The product was extracted with EtOAc (5×20 mL), dried over Na_2SO_4 yielding **23** (47 mg, 70%) as a white foam: ^1H -NMR (300 MHz, $\text{DMSO}-d_6$) δ /ppm: 12.07–11.84 (m, 2H, NH, COOH), 11.32 (s, 1H, NH), 9.32 (s, 1H, NH), 8.57 (s, 1H, NH), 8.30 (t, $J = 5.3$ Hz, 1H, NH), 8.06 (t, $J = 5.4$ Hz, 1H, NH), 6.84–6.72 (m, 3H, NH, $2 \times \text{CH-Pyrr}$), 3.97–3.89 (m, 1H, $\underline{\text{CHCH}_2}$), 3.84–3.65 (m, 2H, COCH_2NH), 3.24–3.17 (m, 2H, $\underline{\text{CH}_2\text{NH}}$), 1.68–1.23 (m, 24H, $3 \times \text{CH}_2$, $2 \times \text{C}(\text{CH}_3)_3$); ^{13}C -NMR (75 MHz, $\text{DMSO}-d_6$) δ /ppm: 172.5 ($\text{C}=\text{O}$), 171.1 ($\text{C}=\text{O}$), 159.4, 159.2, 158.4, 155.3 ($\text{C}=\text{O}$, Boc), 129.7, 125.2, 112.3 (CH-Pyrr), 111.6 (CH-Pyrr), 77.9 ($\underline{\text{C}}(\text{CH}_3)_3$), 54.3 ($\underline{\text{CHCH}_2}$), 40.6 (COCH_2NH), 38.5 ($\underline{\text{CH}_2\text{NH}}$), 31.7 ($\underline{\text{CH}_2\text{CH}}$), 28.8 (CH_2), 28.2 ($\text{C}(\underline{\text{CH}_3})_3$), 27.8 ($\text{C}(\underline{\text{CH}_3})_3$), 22.9 (CH_2). ESI-MS: m/z calcd. for $\text{C}_{25}\text{H}_{38}\text{N}_7\text{O}_9$ $[\text{M} - \text{H}]^-$ 580.27, found 580.1.

Pyrr-C Tripeptide GCP conjugate 24: Compound **1** (32 mg, 0.061 mmol) and GCP-dipeptide **23** (35.55 mg, 0.061 mmol) were dissolved in dry CH_3CN (7 mL) under argon and HOBt (8.27 mg, 0.061 mmol), HBTU (23.20 mg, 0.061 mmol) and dry Et_3N (34 μL , 0.245 mmol) were added. Reaction was stirred at room temperature overnight. Product **24** (28 mg, 47%) was isolated by preparative chromatography ($\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ 9:1) as a white powder: ^1H -NMR (300 MHz, $\text{DMSO}-d_6$) δ /ppm: 11.03 (br. s, 2H, NH, NH-7), 9.59 (br. s, 1H, NH), 9.29 (br. s, 1H, NH), 8.58–8.50 (m, 2H, $2 \times \text{NH}$), 8.34–8.26 (m, 2H, NH, H-4), 8.04 (br. s, 1H, NH), 7.81 (s, 4H, Pht), 6.95–6.73 (m, 3H, CH-Pyrr , $2 \times \text{NH}$), 6.45 (d, $J = 3.7$ Hz, 1H, CH-Pyrr), 5.87 (s, 1H, H-5), 4.64–4.57 (m, 1H, $\underline{\text{CHCH}_2}$), 4.20–4.08 (m, 3H, $\underline{\text{CH}_2\text{Pht}}$, $\underline{\text{CHCH}_2}$), 3.95–3.89 (m, 2H, $\underline{\text{CH}_2\text{Pht}}$), 3.71–3.67 (m, 2H, COCH_2NH), 3.58 (s, 3H, OCH_3), 3.03–2.69 (m, 4H, $\underline{\text{CH}_2\text{CH}}$, $\underline{\text{CH}_2\text{NH}}$), 1.66–1.23 (m, 24H, $3 \times \text{CH}_2$, $2 \times \text{C}(\text{CH}_3)_3$); ^{13}C -NMR (75 MHz, $\text{DMSO}-d_6$) δ /ppm: 172.4 ($\text{C}=\text{O}$), 171.4 ($\text{C}=\text{O}$), 168.9 ($\text{C}=\text{O}$), 168.8, 167.7 ($\text{C}=\text{O}$), 159.5, 158.9 ($\text{C}-2$), 158.6, 158.5, 155.4 ($\text{C}=\text{O}$, Boc), 154.8 ($\text{C}-6$), 140.4 ($\text{C}-4$), 136.8 ($\text{C}-7\text{a}$), 134.3 (Pht), 131.6 (Pht), 123.1 (Pht), 113.6 (CH-Pyrr), 111.7 (CH-Pyrr), 108.6 ($\text{C}-4\text{a}$), 98.2 ($\text{C}-5$), 78.1 ($\underline{\text{C}}(\text{CH}_3)_3$), 63.8 (CH_2), 54.3 ($\underline{\text{CHCH}_2}$), 52.0 (OCH_3), 51.2 (COCH_2NH), 49.3 ($\underline{\text{CH}_2\text{Pht}}$), 41.6 (COCH_2NH), 40.6 ($\underline{\text{CH}_2\text{NH}}$), 37.0 ($\underline{\text{CH}_2\text{Pht}}$), 31.5 ($\underline{\text{CH}_2\text{CH}}$), 29.6 (CH_2), 29.0 (CH_2), 28.2 ($\text{C}(\underline{\text{CH}_3})_3$), 27.7 ($\text{C}(\underline{\text{CH}_3})_3$), 22.8 (CH_2). ESI-MS: m/z calcd. for $\text{C}_{45}\text{H}_{57}\text{N}_{12}\text{O}_{13}$ $[\text{M} + \text{H}]^+$ 973.42, found 973.4.

Pyrr-C Tripeptide GCP conjugate 3: Compound **24** (19 mg, 0.02 mmol) was dissolved in 1:1 mixture of TFA/ CH_2Cl_2 (4 mL) and stirred at room temperature for 20 h. After removal of remaining TFA under reduced pressure, the residue was purified by preparative chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$ 9:1) and product **3** (19 mg, 95%) was obtained as a yellow foam; ^1H -NMR (300 MHz, $\text{DMSO}-d_6$) δ /ppm: 12.20 (br. s, 1H, NH), 11.18–10.97 (m, 2H, NH-7, NH), 8.66 (br. s, 1H, NH), 8.55–8.11 (m, 9H, $2 \times \text{NH}$, $2 \times \text{NH}_3$, H-4), 7.81 (s, 4H, Pht), 7.54 (s, 1H, NH), 7.02 (br. s, 1H, CH-Pyrr), 6.53 (s, 1H, CH-Pyrr), 5.88 (s, 1H, H-5), 4.63 (br. s, 1H, $\underline{\text{CHCH}_2}$), 4.17–4.14 (m, 3H, $\underline{\text{CH}_2\text{Pht}}$, $\underline{\text{CHCH}_2}$), 3.95–3.91 (m, 2H, $\underline{\text{CH}_2\text{Pht}}$), 3.81–3.79 (m, 2H, COCH_2NH), 3.59 (m, 3H, OCH_3), 2.95–2.82 (m, 4H, $\underline{\text{CH}_2\text{CH}}$, $\underline{\text{CH}_2\text{NH}}$), 1.73–1.24 (m, 6H, $3 \times \text{CH}_2$); ^{13}C -NMR (75 MHz, $\text{DMSO}-d_6$) δ /ppm: 169.7 ($\text{C}=\text{O}$), 168.9, 168.2, 167.7, 158.1, 157.6, 154.7 ($\text{C}-6$), 140.4 ($\text{C}-4$), 136.8 ($\text{C}-7\text{a}$), 134.3 (Pht), 131.6 (Pht), 123.0 (Pht), 113.8 (CH-Pyrr), 113.0 (CH-Pyrr), 108.6 ($\text{C}-4\text{a}$), 98.2 ($\text{C}-5$), 62.6 (CH_2), 52.3 ($\underline{\text{CHCH}_2}$), 52.1 (OCH_3), 51.2 (COCH_2NH), 49.3 ($\underline{\text{CH}_2\text{Pht}}$), 41.5 (COCH_2NH), 38.4 ($\underline{\text{CH}_2\text{NH}}$), 37.0 ($\underline{\text{CH}_2\text{Pht}}$), 30.8 ($\underline{\text{CH}_2\text{CH}}$), 28.9 (CH_2), 27.2 (CH_2), 21.4 (CH_2). HRMS (MALDI-TOF/TOF): m/z calcd. for $\text{C}_{35}\text{H}_{40}\text{N}_{12}\text{O}_9$ $[\text{M} + \text{H}]^+$ 773.3114; found 773.3102.

3.3. Computational Details

In order to sample the conformational flexibility of investigated systems and probe their intrinsic dynamical features in the aqueous solution, classical molecular dynamics (MD) simulations were performed employing standard generalized AMBER force fields (ff14SB and GAFF) [51] as

implemented within the AMBER16 program package [52]. All structures were subsequently solvated in a truncated octahedral box of TIP3P water molecules spanning a 10 Å thick buffer of solvent molecules around each system. Upon gradual heating from 0 K, MD simulations were performed at 400 K for a period of 100 ns, maintaining the temperature constant using the Langevin thermostat with a collision frequency of 1 ps^{−1}. Following the clustering analysis of the structures in the obtained MD trajectories, around 20–25 different structures in each case were subjected to geometry re-optimization in Gaussian09 [53] at the M06–2X/6–31G(d) level of theory with aqueous solution being modelled through the implicit SMD solvation. The idea behind this computational strategy was to investigate whether intrinsic dynamical features of investigated conjugates both affect and can explain their tendency to interact with ss-polynucleotides, which avoids difficulties and inaccuracies associated with the computational prediction of the structure of single-stranded polynucleotides, as very recently emphasized by Jeddi and Saiz [54]. pK_a values were calculated in a relative fashion with the MP2/6–311++G(2df,2pd)//(SMD)/M06–2X/6–31+G(d) model, using $AH^+ + B_{REF} \rightarrow A + B_{REF}H^+$ equation, and employing the following reference bases (B_{REF}): succinimide (pK_a = 3.9) [55] for the O-protonation in **I**, cytosine (pK_a = 4.5) [55] for the N-protonation in **II**, and guanidine (pK_a = 13.6) [55] for the N-protonation in **III**. The choice of such computational setup was prompted by its recent success in evaluating solution-phase geometries, reactivities and pK_a values of a large variety of systems [56–59].

4. Conclusions

Two structural isomers of fluorescent cytosine derivative conjugated with guanidiniocarbonyl-pyrrole (GCP) were prepared, whereby they differ in the position of GCP: **2** being more rigid due to the proximity of GCP to the peptide backbone, while **3** is more flexible with GCP attached at the end of a long aliphatic chain of lysine.

In comparison with previously studied GCP derivatives [21–25], here studied **2**, **3** did not induce thermal stabilisation of ds-DNA/RNA nor showed ICD bands of GCP about 300 nm, which supported random aggregation of compounds along double stranded helix of polynucleotide, based on non-specific electrostatic interactions of positive charge with DNA/RNA backbone and hydrophobic interactions. However, even such non-defined binding was highly sensitive to ds-polynucleotide secondary structure, at pH 7 giving fluorescence increase only for ds-DNA, but not ds-RNA.

Intriguingly, **2** and **3** interacted with single stranded (ss) RNA with similar affinity or in case of **3**/poly G order of magnitude higher affinity than with ds-DNA/RNA. With ss-polynucleotides fluorescence emission of nucleobase-phthalimide fluorophore in **2** and **3** became very selective: at pH 7 **2** revealed highly selective emission increase only for poly A, while emission of **3** was selectively quenched only by addition of poly G. Observed selectivities can only be attributed to the positioning of the GCP unit, whereby more flexible combination in **3** prefers poly G as a Watson-Crick partner for cytosine, while more rigid **2** responds only to the mismatched poly A. Molecular modelling confirmed the important role of the GCP fragment in sensing of ss-polynucleotides in solution and showed that these interactions can be rationalized by intrinsic dynamical features of **2** and **3** themselves. Additionally, the evidence is provided that successive recognition of ss-polynucleotides is influenced by a fine interplay between two factors: (a) the protonation state of the GCP moiety; and (b) the ability of the GCP unit to form hydrogen bonds with the rest of the molecule. In general, protonated GCP groups are slightly more susceptible towards ss-polynucleotides (as in **3**²⁺) unless these monocationic units are stabilized intramolecularly, which then diminishes their sensing ability (as in **2**²⁺).

At pH 5 the fluorimetric response of the cytosine-phthalimide fluorophore changed significantly. The poly A selectivity of **2** was completely lost, which could be attributed either to the rearrangement of the polynucleotide into the ds-form poly AH⁺-poly AH⁺ or according to a molecular modelling study (Figure 6), to the stronger intramolecular packing of **2**, in which GCP unit is strongly involved. However, same poly AH⁺-poly AH⁺ induced the fluorescence increase of more flexible **3**, which has

less intramolecularly blocked GCP unit and therefore was able to adapt to a new binding motif within ds-RNA and direct nucleobase fluorophore in interactions causing strong fluorescence change.

The general conclusion is that due to the fine differences in intramolecular interactions of **2** or **3** and their interactions with ss-RNA at different pH it is possible to finely tune recognition of adenine versus guanine sequences, while cytosine and uracil did not cause measurable fluorimetric changes.

Here presented results demonstrated for the first time that the intriguing GCP unit can be applied for the recognition of single stranded RNA sequences—an intuitively unexpected result since so far binding of GCP unit to ds-DNA or ds-RNA relied strongly on minor or major groove interactions. Since ss-polynucleotides do not have structurally defined grooves, clearly GCP unit can take part in H-bonding and/or electrostatic interactions in a selective manner, thus expanding its uses in the design of novel DNA or RNA targeting small molecules. Thus, nucleobase–GCP conjugates can be considered as novel lead compounds for the design of ss-RNA or ss-DNA selective fluorimetric probes.

We have performed the preliminary screening of in vitro activity of **2** and **3** against the human tumor cell lines “HeLa Kyoto” and “HEK293T” by standard MTT test. Cytotoxicity of compounds was tested in the $c = 10^{-4}$ – 10^{-6} M range and no effect on cell proliferation was observed nor cell morphology was changed. These results indicate negligible cytotoxicity of studied compounds and encourage further studies of compounds as in vitro markers.

Supplementary Materials: Supplementary Materials are available online, Additional characterization data, DNA/RNA binding data.

Acknowledgments: Financial support from Croatian Science Foundation (Grant HRZZ-1477) is gratefully acknowledged.

Author Contributions: I.P. and C.S. designed the research concept and wrote the paper; Ž.B. and B.Ž. synthesized new compounds; Ž.B. performed titration experiments; R.V. performed molecular modeling.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Silverman, R.B. *The Organic Chemistry of Drug Design and Drug Action*; Elsevier Academic Press: New York, NY, USA, 2004.
2. Demeunynck, M.; Bailly, C.; Wilson, W.D. *Small Molecule DNA and RNA Binders: From Synthesis to Nucleic Acid Complexes*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2004.
3. Svitkin, Y.V.; Sonenberg, N. An Efficient System for Cap- and Poly(A)-Dependent Translation In Vitro. *Methods Mol. Biol.* **2004**, *257*, 155–170. [[CrossRef](#)]
4. Colgan, D.F.; Murthy, K.G.; Prives, C.; Manley, J.L. Cell-cycle related regulation of poly(A) polymerase by phosphorylation. *Nature* **1996**, *384*, 282–285. [[CrossRef](#)] [[PubMed](#)]
5. Alvarez, E.; Castello, A.; Menendez-Arias, L.; Carrasco, L. HIV protease cleaves poly(A)-binding protein. *Biochem. J.* **2006**, *396*, 219–226. [[CrossRef](#)] [[PubMed](#)]
6. Leroy, J.L.; Gueron, M.; Mergny, J.L.; Helene, C. Intramolecular Folding of a Fragment of the Cytosine-Rich Strand of Telomeric DNA into an I-Motif. *Nucleic Acids Res.* **1994**, *22*, 1600–1606. [[CrossRef](#)] [[PubMed](#)]
7. Manzini, G.; Yathindra, N.; Xodo, L.E. Evidence for intramolecularly folded I-DNA structures in biologically relevant CCC-repeat sequences. *Nucleic Acids Res.* **1994**, *22*, 4634–4640. [[CrossRef](#)] [[PubMed](#)]
8. Giri, P.; Hossain, M.; Kumar, G.S. RNA specific molecules: Cytotoxic plant alkaloid palmatine binds strongly to poly(A). *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2364–2368. [[CrossRef](#)] [[PubMed](#)]
9. Xing, F.; Song, G.; Ren, J.; Chaires, J.B.; Qu, X. Molecular recognition of nucleic acids: Coralyne binds strongly to poly(A). *FEBS Lett.* **2005**, *579*, 5035–5039. [[CrossRef](#)] [[PubMed](#)]
10. Xi, H.J.; Gray, D.; Kumar, S.; Arya, D.P. Molecular recognition of single-stranded RNA: Neomycin binding to poly(A). *FEBS Lett.* **2009**, *583*, 2269–2275. [[CrossRef](#)] [[PubMed](#)]
11. Martelli, A.; Constant, J.F.; Demeunynck, M.; Lhomme, J.; Dumy, P. Design of site specific DNA damaging agents for generation of multiply damaged sites. *Tetrahedron* **2002**, *58*, 4291–4298. [[CrossRef](#)]
12. Lhomme, J.; Constant, J.F.; Demeunynck, M. Abasic DNA structure, reactivity, and recognition. *Biopolymers* **1999**, *52*, 65–83. [[CrossRef](#)]

13. Tumir, L.-M.; Piantanida, I.; Novak, P.; Žinić, M. Interactions of novel phenanthridinium-nucleobase conjugates with complementary and non-complementary nucleotides in aqueous media. *J. Phys. Org. Chem.* **2002**, *15*, 599–607. [[CrossRef](#)]
14. Tumir, L.M.; Piantanida, I.; Juranović Cindrić, I.; Hrenar, T.; Meić, Z.; Žinić, M. New permanently charged phenanthridinium-nucleobase conjugates. Interactions with nucleotides and polynucleotides and recognition of ds-polyAH(+). *J. Phys. Org. Chem.* **2003**, *16*, 891–899. [[CrossRef](#)]
15. Tumir, L.-M.; Piantanida, I.; Žinić, M.; Juranović Cindrić, I.; Meić, Z.; Kralj, M.; Tomić, S. Synthesis of phenanthridinium-bis-nucleobase conjugates, interactions with poly U, nucleotides and in vitro antitumour activity of mono- and bis-nucleobase conjugates. *Eur. J. Med. Chem.* **2006**, *41*, 1153–1166. [[CrossRef](#)] [[PubMed](#)]
16. Juranović, I.; Meić, Z.; Piantanida, I.; Tumir, L.-M.; Žinić, M. Interactions of phenanthridinium-nucleobase conjugates with polynucleotides in aqueous media. Recognition of poly U. *Chem. Commun.* **2002**, 1432–1433. [[CrossRef](#)]
17. Tumir, L.-M.; Piantanida, I.; Juranović Cindrić, I.; Meić, Z.; Tomić, S.; Žinić, M. Recognition of homo-polynucleotides containing adenine by a phenanthridinium bis-uracil conjugate in aqueous media. *Chem. Commun.* **2005**, 2561–2563. [[CrossRef](#)] [[PubMed](#)]
18. Tumir, L.-M.; Grabar, M.; Tomić, S.; Piantanida, I. The interactions of bis-phenanthridinium-nucleobase conjugates with nucleotides: Adenine-conjugate recognizes UMP in aqueous medium. *Tetrahedron* **2010**, *6*, 2501–2513. [[CrossRef](#)]
19. Grabar Branilović, M.; Tomić, S.; Tumir, L.-M.; Piantanida, I. The bis-phenanthridinium system flexibility and position of covalently bound uracil finely tunes the interaction with polynucleotides. *Mol. BioSyst.* **2013**, *9*, 2051–2062. [[CrossRef](#)] [[PubMed](#)]
20. Radić Stojković, M.; Skugor, M.; Tomić, S.; Grabar, M.; Smrečki, V.; Dudek, L.; Grolik, J.; Eilmes, J.; Piantanida, I. Dibenzo-tetraaza[14]annulene-adenine conjugate recognizes complementary poly dT among ss-DNA/ss-RNA sequences. *Org. Biomol. Chem.* **2013**, *11*, 4077–4085. [[CrossRef](#)] [[PubMed](#)]
21. Wu, J.; Zou, Y.; Li, C.; Sicking, W.; Piantanida, I.; Yi, T.; Schmuck, C. A Molecular Peptide Beacon for the Ratiometric Sensing of Nucleic Acids. *J. Am. Chem. Soc.* **2012**, *134*, 1958–1961. [[CrossRef](#)] [[PubMed](#)]
22. Hernandez-Folgado, L.; Baretić, D.; Piantanida, I.; Marjanović, M.; Kralj, M.; Rehm, T.; Schmuck, C. Guanidiniocarbonylpyrrole-Aryl Derivatives: Structure Tuning for Spectrophotometric Recognition of Specific DNA and RNA Sequences and for Antiproliferative Activity. *Chem. Eur. J.* **2010**, *16*, 3036–3056. [[CrossRef](#)] [[PubMed](#)]
23. Gröger, K.; Baretić, D.; Piantanida, I.; Marjanović, M.; Kralj, M.; Grabar, M.; Tomić, S.; Schmuck, C. Guanidiniocarbonyl-pyrrole-aryl conjugates as nucleic acid sensors: Switch of binding mode and spectroscopic responses by introducing additional binding sites into the linker. *Org. Biomol. Chem.* **2011**, *9*, 198–209. [[CrossRef](#)] [[PubMed](#)]
24. Klemm, K.; Radić Stojković, M.; Horvat, G.; Tomišić, V.; Piantanida, I.; Schmuck, C. Interactions of Multicationic Bis(guanidiniocarbonylpyrrole) Receptors with Double-Stranded Nucleic Acids: Syntheses, Binding Studies, and Atomic Force Microscopy Imaging. *Chem. Eur. J.* **2012**, *18*, 1352–1363. [[CrossRef](#)] [[PubMed](#)]
25. Radić Stojković, M.; Piotrowski, P.; Schmuck, C.; Piantanida, I. A short, rigid linker between pyrene and guanidiniocarbonyl-pyrrole induced a new set of spectroscopic responses to the ds-DNA secondary structure. *Org. Biomol. Chem.* **2015**, *13*, 1629–1633. [[CrossRef](#)] [[PubMed](#)]
26. Li, M.; Matković, M.; Piantanida, I.; Schmuck, C. Incorporation of arginine mimetic residue into peptides for recognition of double stranded nucleic acid structure: Binding and aggregation studies. *Bioorg. Med. Chem.* **2017**, *25*, 1875–1880. [[CrossRef](#)] [[PubMed](#)]
27. Hudson, R.H.E.; Dambenieks, A.K.; Viirre, R.D. Fluorescent 7-deazapurine derivatives from 5-iodocytosine via a tandem cross-coupling-annulation reaction with terminal alkynes. *Synlett* **2004**, *13*, 2400–2402. [[CrossRef](#)]
28. Woo, J.S.; Meyer, R.B.; Gamper, H.B. G/C-modified oligodeoxynucleotides with selective complementarity: Synthesis and hybridization properties. *Nucleic Acids Res.* **1996**, *24*, 2470–2475. [[CrossRef](#)] [[PubMed](#)]
29. Liu, C.; Martin, C.T. Fluorescence characterization of the transcription bubble in elongation complexes of T7 RNA polymerase. *J. Mol. Biol.* **2001**, *308*, 465–475. [[CrossRef](#)] [[PubMed](#)]

30. Hudson, R.H.E.; Viirre, R.D.; McCourt, N.; Tse, J. Nucleobase modified peptide nucleic acid. *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 1029–1033. [[CrossRef](#)] [[PubMed](#)]
31. Larock, R.C. Palladium-catalyzed annulation. *J. Organomet. Chem.* **1999**, *576*, 111–124. [[CrossRef](#)]
32. Chang, P.K. 2'-Deoxy-5-iodocytidine 5'-triphosphatetriphosphate and T-deoxy-5-iodouridine 5'-triphosphate. In *Nucleic Acid Chemistry: Improved and New Synthetic Procedures, Methods and Techniques*; Townsend, L.B., Tipson, R.S., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 1978; Volume 2, pp. 779–782.
33. Watanabe, K.A.; Su, T.-L.; Klein, R.S.; Chu, C.K.; Matsuda, A.; Chun, M.W.; Lopez, C.; Fox, J.J. Nucleosides. 123. Synthesis of Antiviral Nucleosides-5-Substituted 1-(2-Deoxy-2-Halogeno-Beta-D-Arabinofuranosyl)Cytosines and -Uracils. Some Structure Activity Relationships. *J. Med. Chem.* **1983**, *26*, 152–156. [[CrossRef](#)] [[PubMed](#)]
34. Holý, A.; Rosenberg, I.; Dvořáková, H. Synthesis of N-(2-Phosphonylmethoxyethyl) Derivatives of Heterocyclic Bases. *Collect. Czechoslov. Chem. Commun.* **1989**, *54*, 2190–2210. [[CrossRef](#)]
35. Hudson, R.H.E.; Dambenieks, A.K. Synthesis of N1-unsubstituted 5-alkynylcytosine and derivatives thereof. *Heterocycles* **2006**, *68*, 1325–1328. [[CrossRef](#)]
36. Dondoni, A.; Perrone, D. Synthesis of 1,1-Dimethylethyl(S)-4-formyl-2,2-dimethyl-3-oxazolidinecarboxylate by oxidation of the alcohol. *Org. Synth.* **2000**, *77*, 64–77. [[CrossRef](#)]
37. Schmuck, C.; Bickert, V.; Merschky, M.; Geiger, L.; Rupprecht, D.; Dudaczek, J.; Wich, P.; Rehm, T.; Machon, U. A facile and efficient multi-gram synthesis of N-protected 5-(Guanidinocarbonyl)-1H-pyrrole-2-carboxylic acids. *Eur. J. Org. Chem.* **2008**, *2*, 324–329. [[CrossRef](#)]
38. Dourtoglou, V.; Gross, B.; Lambropoulou, V.; Zioudrou, C. O-Benzotriazolyl-N,N,N',N'-Tetramethyluronium Hexafluorophosphate as Coupling Reagent for the Synthesis of Peptides of Biological Interest. *Synthesis* **1984**, *7*, 572–574. [[CrossRef](#)]
39. Cantor, C.R.; Scimmell, P.R. *Biophysical Chemistry*; WH Freeman and Co.: San Francisco, CA, USA, 1980; Volume 3, pp. 1109–1181.
40. Scatchard, G. The attractions of proteins for small molecules and ions. *Ann. N.Y. Acad. Sci.* **1949**, *51*, 660–672. [[CrossRef](#)]
41. McGhee, J.D.; Hippel, P.H. Theoretical aspects of DNA-protein interactions: Co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* **1974**, *86*, 469–489. [[CrossRef](#)]
42. Rodger, A.; Norden, B. *Circular Dichroism and Linear Dichroism*; Oxford University Press: New York, NY, USA, 1997; Chapter 2.
43. Eriksson, M.; Nordén, B. Linear and Circular Dichroism of Drug-Nucleic Acid Complexes. *Methods Enzymol.* **2001**, *340*, 68–98. [[CrossRef](#)] [[PubMed](#)]
44. Borštnar, R.; Repič, M.; Kamerlin, S.C.L.; Vianello, R.; Mavri, J. Computational study of the pKa values of potential catalytic residues in the active site of monoamine oxidase B. *J. Chem. Theory Comput.* **2012**, *8*, 3864–3870. [[CrossRef](#)] [[PubMed](#)]
45. Schwamm, R.J.; Vianello, R.; Maršavelski, A.; García, M.A.; Claramunt, R.M.; Alkorta, I.; Saame, J.; Leito, I.; Fitchett, C.M.; Edwards, A.J.; et al. ¹⁵N NMR spectroscopy, X-ray and neutron diffraction, quantum-chemical calculations, and UV/Vis-spectrophotometric titrations as complementary techniques for the analysis of pyridine-supported bicyclic guanidine superbases. *J. Org. Chem.* **2016**, *81*, 7612–7625. [[CrossRef](#)] [[PubMed](#)]
46. Coles, M.P.; Aragon-Saez, P.J.; Oakley, S.H.; Hitchcock, P.B.; Davidson, M.G.; Maksić, Z.B.; Vianello, R.; Leito, I.; Kaljurand, I.; Apperley, D.C. Superbasicity of a bis-guanidino compound with a flexible linker: A theoretical and experimental study. *J. Am. Chem. Soc.* **2009**, *131*, 16858–16868. [[CrossRef](#)] [[PubMed](#)]
47. Lõkov, M.; Tshepelevitsh, S.; Heering, A.; Plieger, P.G.; Vianello, R.; Leito, I. On the basicity of conjugated nitrogen heterocycles in different media. *Eur. J. Org. Chem.* **2017**, *30*, 4475–4489. [[CrossRef](#)]
48. Chaires, J.B.; Dattagupta, N.; Crothers, D.M. Studies on interaction of anthracycline antibiotics and deoxyribonucleic acid: Equilibrium binding studies on interaction of daunomycin with deoxyribonucleic acid. *Biochemistry* **1982**, *21*, 3933–3940. [[CrossRef](#)] [[PubMed](#)]
49. Palm, B.S.; Piantanida, I.; Žinić, M.; Schneider, H.J. The interaction of new 4,9-diazapyrenium compounds with double stranded nucleic acids. *J. Chem. Soc. Perkin Trans.* **2000**, *2*, 385–392. [[CrossRef](#)]
50. Mergny, J.-L.; Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **2003**, *13*, 515–537. [[CrossRef](#)] [[PubMed](#)]

51. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)] [[PubMed](#)]
52. Case, D.A.; Betz, R.M.; Botello-Smith, W.; Cerutti, D.S.; Cheatham, T.E.; Darden, T.A.; Duke, R.E.; Giese, T.J.; Gohlke, H.; Goetz, A.W.; et al. *AMBER 2016*; University of California: San Francisco, CA, USA, 2016.
53. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; et al. *Gaussian 09, Revision A.02*; Gaussian Inc.: Wallingford, UK, 2009.
54. Jeddi, I.; Saiz, L. Three-dimensional modeling of single stranded DNA hairpins for aptamer-based biosensors. *Sci. Rep.* **2017**, *7*, 1178. [[CrossRef](#)] [[PubMed](#)]
55. Cox, B.G. *Acids and Bases: Solvent Effects on Acid-Base Strength*; Oxford University Press: Oxford, UK, 2013, ISBN 978-0199670529.
56. Horak, E.; Vianello, R.; Hranjec, M.; Krištafor, S.; Karminski Zamola, G.; Murković Steinberg, I. Benzimidazole acrylonitriles as multifunctional push-pull chromophores: Spectral characterisation, protonation equilibria and nanoaggregation in aqueous solution. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2017**, *178*, 225–233. [[CrossRef](#)] [[PubMed](#)]
57. Maršavelski, A.; Vianello, R. What a difference a methyl group makes: The selectivity of monoamine oxidase B towards histamine and *N*-methylhistamine. *Chem. Eur. J.* **2017**, *23*, 2915–2925. [[CrossRef](#)] [[PubMed](#)]
58. Hranjec, M.; Horak, E.; Babić, D.; Plavljanić, S.; Srdović, Z.; Murković Steinberg, I.; Vianello, R.; Perin, N. Fluorescent benzimidazo[1,2-*a*]quinolines: Synthesis, spectroscopic and computational studies of protonation equilibria and metal ion sensitivity. *New J. Chem.* **2017**, *41*, 358–371. [[CrossRef](#)]
59. Despotović, I.; Vianello, R. Engineering exceptionally strong oxygen superbases with 1,8-diazanaphthalene di-*N*-oxides. *Chem. Commun.* **2014**, *50*, 10941–10944. [[CrossRef](#)] [[PubMed](#)]

Sample Availability: Samples of the compounds 1–3 are available from the authors.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).