

Novel DNA / RNA-targeting amino acid beacon for the versatile incorporation at any position within the peptide backbone

Tamara Šmidlehner, Ivo Piantanida

Abstract

ONE-POT TANDEM SYNTHESIS WAS FOR THE FIRST TIME APPLIED TO ATTACH FLUOROPHORE TO THE AMINO ACID SIDE CHAIN, YIELDING AMINO ACID READY FOR PEPTIDE COUPLING AT N-TERMINUS, AND ALSO UPON ACTIVATION AT C-TERMINUS. TWO NEW COMPOUNDS DIFFERING ONLY IN FLUOROPHORE-LINKER LENGTH, SHOWED EXCEPTIONAL FLUORIMETRIC AND CD RECOGNITION BETWEEN DS-RNA AND DS-DNA, THUS BEING PROMISING BEACONS FOR VERSATILE PEPTIDE INCORPORATION.

Keywords Fluorescent amino acid · One-pot synthesis · DNA/RNA recognition · fluorescence · CD

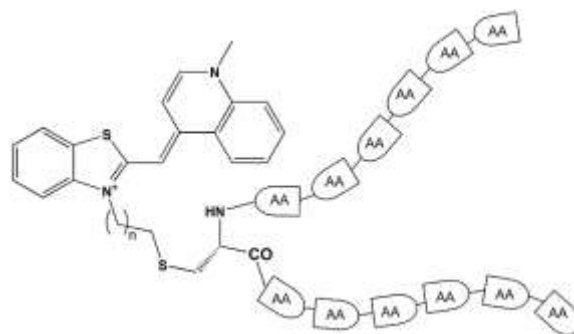
Introduction

Many peptides target various DNA or RNA sequences, controlling a large number of processes in life (Silverman 2004). Among these naturally occurring or synthetic small peptides, many include condensed aryl moiety interacting with DNA/RNA (e.g. intercalator or groove binder), essential for biological activity (Matić et al. 2016). Often aryl-fluorophores are attached to amino acids as fluorescent labels, allowing a huge variety of applications in bioimaging; for instance molecular beacon systems with oligonucleotides or peptide-based molecular beacons (Kummer et al. 2011); (Vasilyeva et al. 2011); (Oh et al. 2009); (Pazos et al. 2009). Only few peptide beacons targeting DNA or RNA have been reported, their function based on Forced Intercalation (FIT) principle (Socher et al. 2008) or due to the conformational change of peptide upon DNA/RNA binding (Wu et al. 2012). Mostly, synthetic approaches are based on attaching the fluorophore to the C- or N- terminus of the peptide, while very few incorporate fluorophore at the amino acid side chain (Socher et al. 2008), although those do

Electronic supplementary material The online version of this article (doi:) contains supplementary material, which is available to authorized users.

T. Šmidlehner, I. Piantanida
Division of Organic Chemistry and Biochemistry,
Ruder Bošković Institute,
Bijenička cesta 54, 10000 Zagreb, Croatia
e-mail: pianta@irb.hr

not use natural peptide bond (O=C–NH). However, only those amino acids which could form natural peptide bonds could be eventually applied as substrates in non-standard amino acid incorporation into proteins (Hong et al. 2014). Additionally, it is quite complicated to prepare peptides with multiple dye-chromophores, which are increasingly interesting for various spectrometric responses upon DNA/RNA binding; for instance for sensing G-quadruplex formation (Meguellati et al. 2010). All aforementioned arguments inspired us to prepare various simple new amino acids with a fluorophore attached to the sidechain, with C- and N- terminus available for incorporation at any position of the peptide by a natural peptide bond. So far, we have shown that fine tuning in distance of peptide-bis-fluorophores can significantly influence spectrometric response (Duksi et al. 2010) and moreover, fluorophore-nucleobase conjugates prepared accordingly demonstrated not only ease of access but also excimer formation dependent on the finely tuned fluorophore-base distance (Duksi et al. 2012). In this work, as a fluorophore, we have chosen well-known cyanine dye TO, and applied the innovative one-pot tandem synthetic procedure to attach it to the amino acid side chain. The advantage of novel procedure is significantly simplified methodology for preparation of amino acid equipped with “fluorescent beacon” unit in off-state, which will light-on exclusively upon binding to DNA/RNA target. Moreover, this amino acid has the N-terminus directly available for peptide coupling, while C- terminus can easily be activated by standard deprotection. Such fluorescent amino acids could be in future applications easily incorporated at any position in the peptide (Scheme 1).



Scheme 1 Two novel amino acids with a fluorophore **TO** attached to the side chain: **4a** (n=1) and **4b** (n=2). Scheme visualises their future potential applications in incorporation at various positions within the peptide.

The difference between two close analogs is only one -CH₂ group in the linker connecting fluorophore with an amino acid. Novel compounds were tested against various ds-DNA and ds-RNA by a set of commonly applied methods and even at monomer level demonstrated surprisingly complex fluorimetric selectivity.

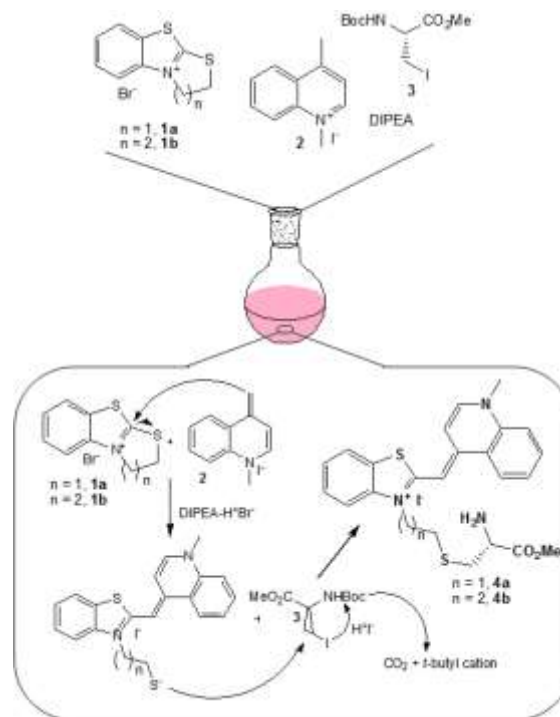
Material and methods

Synthetic procedure

The conventional methods for synthesis of asymmetric monomethine cyanine dyes starting from 2-methylmercaptobenzoxa-(or -thia)-zolium salts (Deligeorgiev et al. 1998); (Hamer 1964); (Deligeorgiev et al. 2006) have several disadvantages: evolution of highly toxic methyl mercaptan and possible position exchange of the alkyl groups at sulfur and nitrogen in the quaternized 2-alkylmercapto starting materials. To avoid the aforementioned disadvantages we used an improved procedure of the recently described method (Japan KokaiTokkyo). Along with avoiding of emission of a strong pollutant, this synthetic approach offers another advantage that provoked our interest- the possibility to generate mercapto-end groups in the formed cyanine dyes, which can be additionally modified as described by Ishiguro et al (1996).

Here presented synthetic procedure applied for the first time aforementioned approach for the preparation of fluorescent amino acid with several novelties and advantages. The most important general advantage is that method is one-pot tandem synthesis consisting of three consecutive reactions (Scheme 2). Namely, leaving group (S⁻ at **1a** or **1b**) from the first reaction acts as a nucleophile in the next step, attacking at the side chain of **3**. At this point, the elimination of HI acidifies the reaction medium and initiates the *in situ* Boc-amino group deprotection (Houben-Weyl 2004), thus yielding product (**4a,b**) with free amino group, conveniently ready for further peptide coupling or in this work ensuring by protonation the additional positive charge in aqueous solution. Advantageous is also very simple workup of the crude product (which has yield over 90%), whereby recrystallization from ether/methanol mixture gives pure compound.

This one-pot procedure enables direct coupling on free amino group with standard coupling reagents, while coupling on C-terminus requires additional reactions (introducing protecting group such as Fmoc on amino group and then removing methyl ester to generate free



Scheme 2 One-pot tandem synthesis procedure of **4a** and **4b**, NOTE immediately available N-terminus for peptide bond formation.

carboxy terminus as proposed by Pascal et al. (1998) suitable for Fmoc-SPPS.

Starting materials **1a** and **1b** were prepared by previously described procedures (Deligeorgiev et al. 2006). Intermediate **2** was prepared by mixing 1 eq. of lepidine and 1.2 eq. of MeI; and amino acid **3** was prepared as described previously (Trost et al. 2003). **1a** or **1b** (0.5 mmol), **2** (0.5 mmol) and **3** (1 mmol) were dissolved in 5 mL of methanol. Then, 0.5 mmol of DIPEA was added and the reaction mixture was stirred overnight at room temperature. Final products were filtered off and washed with ether (conversion on crude product estimated about 90%); recrystallization from methanol giving a pure product of **4a** and **4b** as red solid in 80% yield.

NMR spectra were recorded on 600 and 300 MHz spectrometers, operating at 150.92 or 75.47 MHz for ¹³C and 600.13 or 300.13 MHz for ¹H nuclei. TMS was used as an internal standard. HRMS analysis was performed on MALDI-TOF mass spectrometer operating in reflectron mode. Mass spectra were acquired by accumulating three spectra after 400 laser shots per spectrum. Calibrant and analyte spectra were obtained in positive ion mode. Calibration type was internal with calibrants produced by matrix ionization (monomeric, dimeric and trimeric CHCA), with azithromycin and angiotensin II dissolved in α-cyano-

4-hydroxycinnamic acid matrix in the mass range m/z 190.0499 to 749.5157 or 1046.5417. Accurately measured spectra were internally calibrated and elemental analysis was performed on Data Explorer v. 4.9 Software with mass accuracy better than 5 ppm. Samples were prepared by mixing 1 μL of analyte methanol solution with 5 μL of saturated (10 mg/mL) solution of α -cyano-4-hydroxycinnamic acid (α -CHCA) and internal calibrants (0.1 mg/mL) dissolved in 50 % acetonitrile/0.1 % TFA.

3-(2-{[(2R)-2-amino-3-methoxy-3-oxopropyl]sulfanyl}ethyl)-2-{[(4E)-1-methyl-1,4-dihydroquinolin-4-ylidene]methyl}-1,3-benzothiazol-3-ium iodide (4a). ^1H NMR (300 MHz, DMSO) δ/ppm : 8.81 (d, $J = 8.4$ Hz, 1H, Ar), 8.71 – 8.64 (m, 2H, Ar), 8.13 – 8.01 (m, 4H, Ar + NH_2), 7.80 (d, $J = 8.0$ Hz, 2H, Ar), 7.65 – 7.56 (m, 1H, Ar), 7.48 – 7.41 (m, 2H, Ar), 7.00 (s, 1H, CH), 4.89 (dd, $J = 13.8$, 6.6 Hz, 2H, N^+CH_2), 4.24 – 4.15 (m, 4H, $\text{N}^+\text{CH}_3 + \alpha\text{H}$), 3.65 (s, 3H, CO_2CH_3), 3.17 – 3.05 (m, 4H, $\text{CH}_2\text{-S-CH}_2$); ^{13}C NMR (151 MHz, DMSO) δ/ppm : 159.26 (CO_2CH_3), 148.99 (C_q , Ar), 145.28 (CH, Ar), 139.96 (C_q , Ar), 138.05 (C_q , Ar), 133.31 (CH, Ar), 128.10 (CH, Ar), 127.07 (CH, Ar), 125.67 (CH, Ar), 124.52 (C_q , Ar), 124.15 (C_q , Ar), 123.57 (CH, Ar), 122.89 (CH, Ar), 118.33 (CH, Ar), 113.03 (CH, Ar), 112.99 (CH, Ar), 108.33 ($\text{CH}_3\text{-N}$), 88.00 (CH, Ar), 51.96 (CO_2CH_3), 45.31 ($\text{CH}_2\text{-N}^+$), 42.52 (αC), 29.59 ($\text{CH}_2\text{-S}$), 29.34 (S-CH_2); HRMS (MALDI-TOF/TOF): calcd for $\text{C}_{24}\text{H}_{26}\text{N}_3\text{O}_2\text{S}_2^+ [\text{M}]^+$: 452.1466; found: 452.1480.

3-(2-{[(2R)-2-amino-3-methoxy-3-oxopropyl]sulfanyl}propyl)-2-{[(4E)-1-methyl-1,4-dihydroquinolin-4-ylidene]methyl}-1,3-benzothiazol-3-ium iodide (4b). ^1H NMR (300 MHz, DMSO) δ/ppm : 8.54 (d, $J = 7.1$ Hz, 2H, Ar), 7.94 (t, $J = 7.7$ Hz, 2H, Ar), 7.85 – 7.63 (m, 4H, Ar + NH_2), 7.53 (t, $J = 7.6$ Hz, 2H, Ar), 7.39 (d, $J = 7.3$ Hz, 1H, Ar), 7.13 (d, $J = 7.1$ Hz, 1H, Ar), 6.73 (s, 1H, CH), 4.58 (dd, $J = 12.0$, 6.3 Hz, 2H, N^+CH_2), 4.17 (m, 4H, $\text{N}^+\text{CH}_3 + \alpha\text{H}$), 3.62 – 3.53 (m, 2H, SCH_2), 3.50 (s, 1H, CO_2CH_3), 3.00 (t, $J = 7.1$ Hz, 2H, CH_2S), 2.20 – 2.05 (m, 2H, CH_2); ^{13}C NMR (75 MHz, DMSO) δ/ppm : 158.80 (CO_2CH_3), 148.27 (C_q , Ar), 144.88 (CH, Ar), 139.62 (C_q , Ar), 137.75 (C_q , Ar), 132.96 (CH, Ar), 128.60 (CH, Ar), 127.60 (CH, Ar), s), 126.66 (CH, Ar), 124.94 (C_q , Ar), 124.42 (CH, Ar), 123.72 (CH, Ar), 118.16 (CH, Ar), 114.91 (CH, Ar), 112.56 (CH, Ar), 107.90 ($\text{CH}_3\text{-N}$), 87.02 (CH, Ar), 46.82 ($\text{CH}_2\text{-N}^+$), 42.48 (αC), 28.99 ($\text{CH}_2\text{-S}$), 27.51 (S-CH_2), 19.84 (CH_2).); HRMS (MALDI-TOF/TOF): calcd for $\text{C}_{25}\text{H}_{28}\text{N}_3\text{O}_2\text{S}_2^+ [\text{M}]^+$: 466.1623; found: 466.1601.

Interactions of 4a, 4b with DNA and RNA

All measurements were performed in aqueous buffer solution ($\text{pH} = 7.0$, $I = 0.05$ M, sodium cacodylate/HCl buffer). The UV-Vis spectra were recorded on a Varian Cary 100 Bio spectrometer and fluorescence spectra were recorded on Varian Cary Eclipse fluorimeter in quartz cuvettes (1 cm). Under the experimental conditions used ($0.1 - 1 \cdot 10^{-6}$ M) the absorbance intensities of **4a**, **4b** were proportional to their concentrations (Fig. S1). Polynucleotides were purchased as noted: poly dGdC – poly dGdC, poly dAdT – poly dAdT, poly A – poly U, (Sigma), *calv thymus* (ct)-DNA (Aldrich) and dissolved in sodium cacodylate buffer, $I = 0.05$ M, $\text{pH} = 7.0$. The ct-DNA was additionally sonicated and filtered through a 0.45 mm filter (Chaires et al. 1982). Polynucleotide concentration was determined spectroscopically (Malojčić et al. 2005) as the concentration of phosphates. In fluorimetric experiments, excitation wavelengths at $\lambda_{\text{max}} = 501$ nm were used in order to avoid absorption of excitation light by added polynucleotides. Fluorimetric titrations were performed by adding portions of polynucleotide solution into the solution of the studied compound ($c = 5 \cdot 10^{-7}$ M). After mixing polynucleotides with studied compounds it was observed in all cases that equilibrium was reached in less than 120 seconds. Fluorescence spectra were collected at excess of DNA/RNA $r < 0.1$ ($r = (\text{compound}) / (\text{polynucleotide})$) to assure one dominant binding mode. Data processed by the non-linear fitting procedure to Scatchard equation (Scatchard 1949) gave values of ratio n [bound compound] / [polynucleotide] in the range 0.1-0.3, but for easier comparison, all K_s values were recalculated for the fixed $n = 0.2$. Calculated values for K_s have satisfactory correlation coefficients (>0.99). Relative fluorescence quantum yields (Q) were determined by the standard procedure (Miller 1981); (Gore 2000). All samples were purged with argon to displace oxygen, and emission spectra were recorded in the range from 500-600 nm and corrected for the effects of time- and wavelength-dependent light source fluctuations by use of a rhodamine 101 standard, a diffuser and the software provided with the instrument. As the standard, we used rhodamine 6G (Fluka, Buchs, Switzerland) with published fluorescence quantum yield $Q = 0.95$ (Kubin et al. 1982). The rhodamine 6G concentration ($c = 3.4 \cdot 10^{-7}$ mol dm^{-3}) in fluorescence measurements had an optical absorbance below 0.05 at the excitation wavelength. **4a** and **4b** ($c = 5 \cdot 10^{-7}$ mol dm^{-3}) and their complexes with polynucleotides had an optical absorbance below 0.05 at the excitation wavelength ($\lambda_{\text{exc}} = 499$ nm).

CD spectra were recorded on JASCO J815 spectrophotometer at room temperature using

appropriate 1 cm path quartz cuvettes with a scanning speed of 200 nm/min. Buffer background was subtracted from each spectrum, while each spectrum was the result of three accumulations. CD experiments were performed by adding portions of compound stock solution into the solution of the polynucleotide ($c = 2 \times 10^{-5}$ M). Thermal melting experiments were performed on a Varian Cary 100Bio spectrometer in quartz cuvettes (1 cm). The measurements were done in aqueous buffer solution at pH 7.0 (sodium cacodylate buffer $I = 0.05$ M). Thermal melting curves for ds-DNA, ds-RNA and their complexes with 4a, 4b were determined by following the absorption change at 260 nm as a function of temperature (Mergny et al. 2003). 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. The 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 ± 0.5 °C. High thermal melting of the poly dGdC – poly dGdC (ΔT_m value > 100 °C), didn't allow accurate determination of melting profile, nor study of compound impact on thermal stability of poly dGdC – poly dGdC.

Results

Novel compounds were well soluble in aqueous solutions. Their UV/Vis spectra differed in the shape (ESI Fig. S2), **4a** with a shorter linker having a maximum at 501 nm and **4b** with longer linker at 473 nm. Detailed analysis of concentration dependence of UV/Vis spectra (ESI, Fig. S1, Table S1) revealed that absorption properties of **4a** were proportional to its concentration, while the UV/Vis spectrum of **4b** exhibited pronounced changes: at 0.1 μ M concentrations ratio between 473 and 501 maximum was almost 1:1 but ratio strongly increased with concentration (ESI Fig. S3). Moreover, only UV/Vis spectrum of **4b** significantly changed with temperature increase (ESI, Figure S4.), the ratio between 473 and 501 nm decreasing. Observed results support self-aggregation of **4b** at μ M concentration range, while **4a** in the same conditions is dominantly in monomeric form, most likely due to the shorter linker. The absorption maximum of **4b** monomeric form seems to have a maximum at 501 nm, while aggregated form apparently absorbs at 473 nm, resulting in the significant hypsochromic shift (-30 nm). Such aggregation-induced shift to shorter wavelengths was originally observed also for TO-dye (Nygren et al. 1998) and in accordance with similar changes observed

for merocyanine dyes suggests the formation of H-type aggregate (ESI Figs. S2-4) (Rosch et al. 2006). Both compounds at μ M concentrations exhibited negligible fluorescence emission.

Addition of **4a** and **4b** stabilized moderately all ds-polynucleotides with only exception of **4b**/pApU complex, latter showing negligible thermal destabilization effect (Table 1, ESI Figs S5-10). Mostly stabilization effects were non-linearly dependent on ratio r , suggesting that saturation of dominant binding sites was accomplished about $r = 0.2$ (Mergny et al. 2003); (Wilson et al. 1993). However, the stabilisation effect for the most polynucleotides didn't stop at $r=0.2$ but further increased at $r>0.2$ due to the secondary binding modes of aggregated dyes at excess of dye over DNA ($r>0.2$, see CD results Figure 2), which still contribute to double helix stability. Slightly stronger stabilization effect is observed for p(dAdT)₂ compared to mixed sequence ctDNA (48% of GC-basepairs), which indicated that stabilization is basepair dependent.

Table 1. The ΔT_m^a values (°C) of studied ds-polynucleotides upon addition of **4a**, **4b** at pH 7.0 (sodium cacodylate buffer, $I = 0.05$ M).

r^b	4a		4b			
	ct-DNA	dAdT	pApU	ct-DNA	dAdT	pApU
0.1	1.5	2.8	2.1	1.8	3.1	-0.4
0.2	2.9	4.8	3.9	5.5	6.2	-0.6
0.3	5.8	7.4	4.8	6.1	7.0	-1.0

^a Error in $\Delta T_m \pm 0.5$ °C. ^b $r = [c(\text{compound})] / [c(\text{polynucleotide})]$.

The UV/Vis spectra of new dyes revealed strong hypochromic and moderate bathochromic effect upon titration with ds-DNA and ds-RNA. Detailed analysis of titration data was hampered by aggregation of **4b**, since it was not possible to distinguish the effects of the aggregation-disaggregation equilibrium from the dye monomer binding to DNA and eventually formation of dye aggregate within the DNA/RNA (often reported for cyanine dyes (Crnolatac et al. 2016); (Crnolatac et al. 2013); (Tumir et al. 2012)). Fortunately, intrinsically non-fluorescent **4a**, **4b** upon titration with polynucleotides revealed strong emission (Fig. 1), which allowed experiments to be done at much lower concentrations, at which aggregation of dyes could be neglected. Moreover, aggregates are non-fluorescent and thus the fluorescence emission could be attributed to single dye molecules bound to

DNA/RNA *via* well-established mechanism for cyanine dyes (Armitage 2005).

Strong emission increase allowed accurate determination of binding constants and binding ratios $n_{[\text{bound dye}]/[\text{polynucleotide}]}$ (Table 2). The affinity of **4a**, **4b** toward all studied ds-DNA and RNA were within micromolar range, pointing out toward similar set of dominant binding interactions. At variance to non-selective affinity ($\log K_s$), the fluorescence of **4a**, **4b** was strongly dependent on polynucleotide structure.

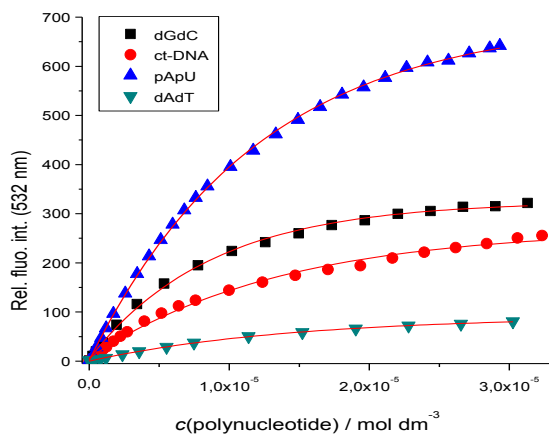


Fig. 1 Induced fluorescence of **4a** ($\lambda_{\text{exc}} = 501 \text{ nm}$, $\lambda_{\text{em}} = 532 \text{ nm}$, $c = 5.2 \times 10^{-7} \text{ M}$) upon addition of polynucleotides, at pH 7.0, Na cacodylate buffer, $I = 0.05 \text{ M}$. Red lines (—) represent non-linear fitting procedure to Scatchard eq. (Scatchard 1949).

Table 2 Binding constants ($\log K_s^a$), relative quantum yields^b of **4a**, **4b** complexes with ds-polynucleotides calculated from fluorimetric titrations (Fig. 1), at pH = 7.0, Na cacodylate buffer, $I = 0.05 \text{ M}$.

		ct-DNA	dAdT	dGdC	pApU
	$\log K_s^a$	5.5	5.4	5.8	5.5
4a	Q^b	0.071	0.014	0.039	0.095
	$\log K_s^a$	5.2	5.5	6-7 ^c	5.0
4b	Q^b	0.075	0.012	0.039	0.076

^a Processing of titration data by non-linear fitting procedure to Scatchard equation (Scatchard 1949) gave values of ratio $n_{[\text{bound dye}]/[\text{polynucleotide}]} = 0.1-0.3$, 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 Relative quantum yield (Q) was determined in respect to the rhodamine 6G standard (Kubin et al. 1982) ($Q = 0.95$). ^c Small changes in emission allowed only estimation of affinity.

The **4a** showed for AU-RNA 7-fold stronger emission intensity in comparison to AT-DNA (Fig. 1), while fluorescence increase of **4b** was only 5-fold higher for AU-RNA than with AT-DNA (Table 2). All aforementioned differences in interactions with DNA/RNA strongly stress not only the impact of amino acid attached to N3-position of thiazole but also the influence of the linker length.

To study in more structural detail binding of **4a**, **4b** to DNA/RNA we used CD spectropolarimetry, a highly informative method for the study of conformational changes in the secondary structure of DNA/RNA (Rodger and Norden 1997). In addition, small molecules can eventually acquire induced CD spectrum (ICD) upon binding to polynucleotides, which could give useful information about modes of interaction (Eriksson and Norden 2001). Although **4a**, **4b** possess intrinsic CD spectra (ESI Fig S11), their intensity under biologically relevant conditions used is negligible. Addition of **4a** and **4b** to all polynucleotides yielded significant changes in CD spectra of polynucleotides ($\lambda < 300 \text{ nm}$) as well as the appearance of strong induced CD spectra (ICD) at $\lambda = 400-600 \text{ nm}$ (Fig. 2, ESI Figs S12-13). At DNA/RNA over dye excess ($r < 0.05$) **4a**, **4b** bind as single molecules into dominant binding sites of DNA or RNA, showing in the dye ICD region ($\lambda = 400-600 \text{ nm}$) single, and very weak ICD bands, characteristic for intercalative binding mode (Rodger and Norden 1997).

However, at higher ratios ($r > 0.1$) in $\lambda = 400-600 \text{ nm}$ region, strong bisignate ICD patterns appeared, typical for dye aggregate (Wilson et al. 1993), whereby isoelliptic points observed for most titrations (Fig. 2, ESI Figs. S12-13) support formation of one dominant type of complex. The shape and sign of bisignate exciton couplets for DNA and RNA is similar (negative/positive) but intriguingly, sets for DNA are significantly bathochromically shifted with respect to RNA. Moreover, only **4b** showed additional negative ICD band for RNA at 540 nm, which matches the positive band of **4b**/DNA complex.

Observed results show that **4a** and **4b** form aggregates in DNA or RNA grooves, whereby organization of aggregate is very sensitive to the binding site environment, which is reported by significant differences in bisignate exciton couplet ICD patterns. According to previous results (Tumir et al. 2012) cyanine-dimers nicely fit within the AT-AT DNA groove (giving typical ICD pattern), while RNA major groove with similar width as AT-DNA but much larger depth, gives a different frame for dye aggregation, resulting in blue shifted ICD bands. The derivative with a longer linker (**4b**) has less motional freedom

within the RNA major groove and thus exhibits additional ICD band at 540 nm.

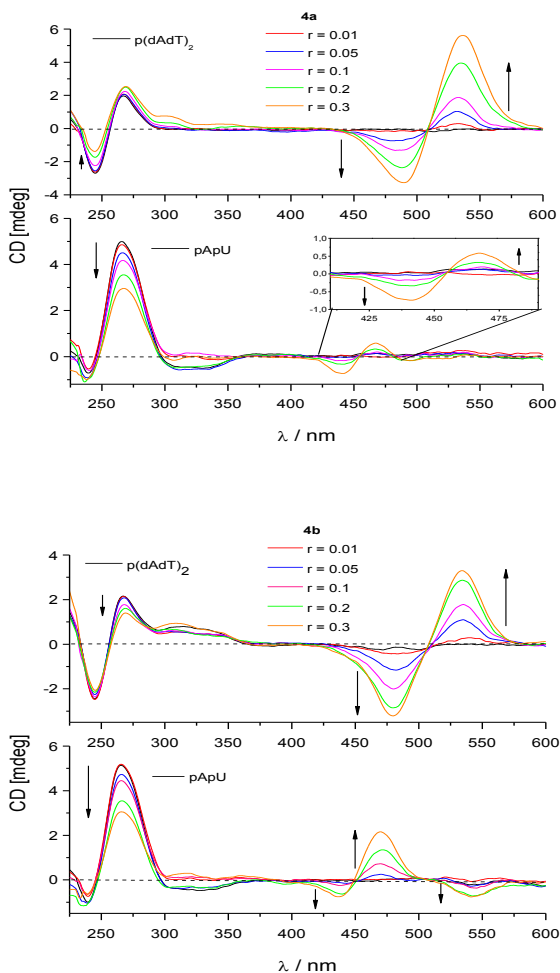


Fig. 2 CD titration of p(dAdT)₂ ($c = 2 \times 10^{-5}$ M) and pApU ($c = 2 \times 10^{-5}$ M) with **4a** (UP) and **4b** (DOWN) at molar ratios $r = [\text{compound}] / [\text{polynucleotide}]$ (pH 7.0, buffer sodium cacodylate, $I = 0.05$ M).

Conclusions

Novel cyanine-amino acid conjugates **4a**, **4b** were synthesized in eco-friendly and simple one-pot reaction. The main advantage of novel methodology is preparation of amino acid equipped with “fluorescent beacon” unit in off-state, which exhibits strong emission exclusively upon binding to DNA/RNA target. Moreover, this amino acid has the N-terminus directly available for peptide coupling, while C-terminus can easily be activated by standard deprotection.

Although the difference between **4a** and **4b** is only one CH₂ group of the linker, their aqueous solutions

significantly differ in aggregation properties; longer **4b** being at 20-micromolar concentrations significantly aggregated, while shorter **4a** mostly present in the monomeric form. Already as mono-amino-acid-dye synthons, **4a**, **4b** revealed bio-applicable binding affinity toward DNA/RNA and even more important, the selective spectrophotometric recognition between DNA and RNA. Compounds **4a**, **4b** bind to DNA and RNA at $r_{[\text{compound}]/[\text{polynucleotide}]} < 0.05$ as intercalators, and additionally at dye-over-DNA/RNA excess ($r > 0.1$) **4a**, **4b** form aggregates within the DNA/RNA grooves. Both compounds at submicromolar concentrations exhibit selective response between ds-DNA and ds-RNA by two methods; fluorescence and CD spectrophotometry. Although **4a**, **4b** are structurally very similar (difference is one CH₂ group in dye-amino acid linker), the selectivity is quite different: the shorter **4a** showed much stronger fluorescence differentiation between AU-RNA and AT-DNA, than it was shown by longer **4b**, and only **4b** (not **4a**) revealed for ds-RNA additional ICD band at 540 nm. These results show that linker length between an amino acid and dye finely tuned interactions with DNA/RNA and thus strongly support future studies of oligopeptides with variously incorporated **4a**, **4b**.

Thus, here presented results demonstrate that novel amino acid-dye conjugates **4a**, **4b** are able to accurately differentiate between ds-DNA and ds-RNA by two highly sensitive methods. That strongly encourages further development of peptides (Scheme 1) in which **4a**, **4b** would be incorporated into any position of short peptides, which is an advantage in respect to current approaches, which usually attach commercial cyanine dye only at the free amino- or carboxyl- groups of the peptide. One approach would be to insert single dye into various positions of peptide targeting particular DNA or RNA sequence/type to see which combination gives the best response/recognition/stabilisation of the target. Even more promising is systematic study peptides incorporating two or more amino acid-dye units at various mutual distances, allowing preorganisation of dye-chromophores prior or after the DNA/RNA binding, similar as noted for sensing G-quadruplex formation (Meguellati et al. 2010).

Acknowledgements

This work has been supported by Croatian Science Foundation project 1477. Special thanks to Prof. Todor Deligeorgiev and dr. Atanas Kurutos for fruitful discussions and advice.

Compliance with ethical standards

Conflict of interest: The authors declare that there are no conflicts of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent

Informed consent was obtained from all individual participants included in the study.

References

- Armitage (2005) DNA Binders and Related Subjects: Cyanine dye-DNA interactions: Intercalation, groove binding, and aggregation. *Top Curr Chem* 253:55–76. doi: 10.1007/B100442
- Chaires JB, Dattagupta N CD (1982) Studies on Interaction of Anthracycline Antibiotics and Deoxyribonucleic-Acid - Equilibrium Binding-Studies on Interaction of Daunomycin with Deoxyribonucleic Acid. *Biochemistry* 21:3933–3940. doi: 10.1021/bi00260a005
- Crnolatac I, Rogan I, Majic B, Tomic S, Deligeorgiev T, Horvat G, Makuc D, Plavec J, Pescitelli G PI (2016) Small molecule probes finely differentiate between various ds- and ss-DNA and RNA by fluorescence, CD and NMR response. *Anal Chim Acta* 940:128–135. doi: 10.1016/j.aca.2016.08.021
- Crnolatac I, Tumir LM, Lesev NY, Vasilev AA, Deligeorgiev TG, Miskovic K, Glavas-Obrovac L, Vugrek O PI (2013) Probing the Structural Properties of DNA/RNA Grooves with Sterically Restricted Phosphonium Dyes: Screening of Dye Cytotoxicity and Uptake. *ChemMedChem* 7:1093–1103. doi: 10.1002/cmdc.201300085
- Deligeorgiev TG, Zaneva D a, Kim SH, Sabnis RW (1998) Preparation of monomethine cyanine dyes for nucleic acid detection. *Dye Pigment* 37:205–211. doi: 10.1016/S0143-7208(97)80088-5
- Deligeorgiev T, Gadjev N, Vasilev A, Drexhage KH YS (2006) Synthesis of novel monomeric cyanine dyes containing mercapto and thioacetyl substituents for nucleic acid detection. *Dye Pigment* 7:185–191. doi: 10.1016/j.dyepig.2005.05.010.
- Duksi M, Baretic D, Caplar V PI (2010) Novel bis-phenanthridine derivatives with easily tunable linkers, study of their interactions with DNA and screening of antiproliferative activity. *Eur J Med Chem* 45:2671–2676. doi: 10.1016/j.ejmech.2010.02.017
- Duksi M, Baretic D PI (2012) Synthesis of the Peptide-Based Phenanthridine-Nucleobase Conjugates and Study of Their Interactions With ds-DNA. *Acta Chim Slov* 59:464–472.
- G. S (1949) The attractions of proteins for small molecules and ions. *Ann NY Acad Sci* 51:660–672.
- Gore MG (2000) Spectrophotometry and Spectrofluorimetry. A practical Approach. Oxford University Press
- Hamer FM (1964) The cyanine dyes and related compounds. Interscience Publisher, New York
- Hong SH, Kwon YC JM (2014) Non-standard amino acid incorporation into proteins using *Escherichia coli* cell-free protein synthesis. *Front Chem* 2:1–7. doi: ARTN 34 10.3389/fchem.2014.00034
- Houben-Weyl (2004) Synthesis of Peptides and Peptidomimetics. In: Murray G, Arthur F, Luis M, Claudio T (eds) *Methods in Organic Chemistry*. Thieme Verlag Stuttgart, Stuttgart,
- Ishiguro T, Saitoh J, Yawata H, Otsuka M, Inoue T SY (1996) Fluorescence detection of specific sequence of nucleic acids by oxazole yellow-linked oligonucleotides. Homogeneous quantitative monitoring of in vitro transcription. *Nucl Acids Res* 24:4992–4997. doi: 10.1093/nar/24.24.4992.
- Japan KokaiTokkyo, Koho JP 58 205 144; Japan KokaiTokkyo KJ 58 203 432 No Title.
- Kubin RF, Fletcher AN (1982) Fluorescence quantum yields of some rhodamine dyes. *J Lumin* 27:455–462.
- Malojčić G, Piantanida I, Marinić M, Žinić M, Marjanović M, Kralj M, Pavelić K SH (2005) A novel bis-phenanthridine triamine with pH controlled binding to nucleotides and nucleic acids. *Org Biomol Chem* 3:4373–4381. doi: 10.1039/B509094f
- Matić J, Lidija T, Marijana RS, Piantanida I (2016) Advances in Peptide-based DNA/RNA-Intercalators. *Curr Protein Pept Sc* 17:127–134. doi: 10.2174/138920371702160209124439
- Meguelliati K, Koripelly G LS (2010) DNA-Templated Synthesis of Trimethine Cyanine Dyes: A Versatile Fluorogenic Reaction for Sensing G-Quadruplex Formation. *Angew Chem Int Ed* 49:2738–2742. doi: 10.1002/anie.201000291
- Mergny JL LL (2003) Analysis of thermal melting curves. *Oligonucleotides* 13:515–537. doi: 10.1089/154545703322860825
- Miller JN (1981) Standards for Fluorescence Spectrometry. Chapman and Hall, London
- Nygren J, Svanvik N KM (1998) The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* 46:39–51. doi: 10.1002/(Sici)1097-0282(199807)46:1<39::Aid-Bip4>3.0.Co;2-Z.
- Oh KJ, Cash KJ PK (2009) Beyond Molecular Beacons: Optical Sensors Based on the Binding-Induced

- Folding of Proteins and Polypeptides. *Chem-Eur J* 15:2244–2251. doi: 10.1002/chem.200701748.
- Pascal R, Sola R (1998) Preservation of the Fmoc protective group under alkaline conditions by using CaC12. Applications in peptide synthesis. *Tetrahedron Lett* 39:5031–5034.
- Pazos E, Vazquez O, Mascarenas JL VM (2009) Peptide-based fluorescent biosensors. *Chem Soc Rev* 38:3348–3359. doi: 10.1039/b908546g.
- R.B. S (2004) *The Organic Chemistry of Drug Design and Drug Action*. Elsevier Academic Press, New York
- Rodger A, Norden B (1997) *Circular Dichroism and Linear Dichroism*. Oxford University Press, New York
- Rosch U, Yao S, Wortmann R WF (2006) Fluorescent H-aggregates of merocyanine dyes. *Angew Chem Int Ed* 45:7026–7030. doi: 10.1002/anie.200602286.
- S. K, A K, E S, et al (2011) Fluorescence Imaging of Influenza H1N1 mRNA in Living Infected Cells Using Single-Chromophore FIT-PNA. *Angew Chem Int Ed* 50:1931–1934. doi: 10.1002/anie.201005902
- Socher E, Jarikote DV, Knoll A, Roglin L, Burmeister J SO (2008) FIT probes: Peptide nucleic acid probes with a fluorescent base surrogate enable real-time DNA quantification and single nucleotide polymorphism discovery. *Anal Biochem* 375:318–330. doi: 10.1016/j.ab.2008.01.009.
- Trost BM RM (2003) Chemoselectivity of the ruthenium-catalyzed hydrative diyne cyclization: Total synthesis of (+)-cylindricine C, D, and E. *Org Lett* 5:4599–4602. doi: 10.1021/ol035752n.
- Tumir LM, Crnolatac I, Deligeorgiev T, Vasilev A, Kaloyanova S, Branilovic MG, Tomic S PI (2012) Kinetic Differentiation between Homo- and Alternating AT DNA by Sterically Restricted Phosphonium Dyes. *Chem - A Eur J* 18:3859–3864. doi: 10.1002/chem.201102968.
- Vasilyeva E, Lam B, Fang ZC, Minden MD, Sargent EH KS (2011) Direct Genetic Analysis of Ten Cancer Cells: Tuning Sensor Structure and Molecular Probe Design for Efficient mRNA Capture. *Angew Chem Int Ed* 50:4137–4141. doi: 10.1002/anie.201006793.
- W. D. Wilson, L. Ratmeyer, M. Zhao LS and DB (1993) The search for structure-specific nucleic acid-interactive drugs: effects of compound structure on RNA versus DNA interaction strength. *Biochemistry* 32:4098–4104.
- Wu JC, Zou Y, Li CY, Sicking W, Piantanida I, Yi T SC (2012) A Molecular Peptide Beacon for the Ratiometric Sensing of Nucleic Acids. *J Am Chem Soc* 134:1958–1961. doi: 10.1021/Ja2103845