

Cite this: *Chem. Sci.*, 2016, 7, 3581

# Violation of DNA neighbor exclusion principle in RNA recognition†

Muhammad Yousuf,<sup>‡ab</sup> Il Seung Youn,<sup>‡ab</sup> Jeonghun Yun,<sup>ab</sup> Lubna Rasheed,<sup>b</sup> Rosendo Valero,<sup>b</sup> Genggongwo Shi<sup>ab</sup> and Kwang S. Kim<sup>\*ab</sup>

DNA intercalation has been very useful for engineering DNA-based functional materials. It is generally expected that the intercalation phenomenon in RNA would be similar to that in DNA. Here we note that the neighbor-exclusion principle is violated in RNA by naphthalene-based cationic probes, in contrast to the fact that it is usually valid in DNA. All the intercalation structures are responsible for the fluorescence, where small naphthalene moieties are intercalated in between bases *via*  $\pi$ - $\pi$  interactions. The structure is aided by hydrogen bonds between the cationic moieties and the ribose-phosphate backbone, which results in specific selectivity for RNA over DNA. This experimentally observed mechanism is supported by computationally reproducing the fluorescence and CD data. MD simulations confirm the unfolding of RNA due to the intercalation of probes. Elucidation of the mechanism of selective sensing for RNA over DNA would be highly beneficial for dynamical observation of RNA which is essential for studying its biological roles.

Received 2nd October 2015  
Accepted 14th February 2016

DOI: 10.1039/c5sc03740a

[www.rsc.org/chemicalscience](http://www.rsc.org/chemicalscience)

## 1. Introduction

The neighbor-exclusion principle is a well-known rule for intercalative binding of small planar molecules to DNA.<sup>1-6</sup> According to this principle, the two neighboring sites of an occupied intercalation site in DNA must remain unoccupied or, in less absolute terms, intercalation is anti-cooperative at adjacent sites.<sup>3,6</sup> Namely, every second (next-neighbor) intercalation site along the length of the DNA double helix remains unoccupied. The concept of neighbor-exclusion was originally postulated in consideration of possible stereochemical constraints imposed by the sugar-phosphodiester backbone, but the effects of vibrational entropy and counterion release favor the flexible neighbor-exclusion models over the rigid neighbor-exclusion-violating models.<sup>6</sup> Such neighbor-exclusion states were frequently noted in DNA systems.<sup>1-6</sup> In this study, we show a clear example that such neighbor exclusion states do not work in RNA systems because of differences in the sugar-phosphodiester backbone between DNA and RNA. Specifically, we tested the unique properties of the naphthalene moiety which are responsible for the violation of the neighbor-exclusion principle in RNA, thereby imparting an effect for highly selective recognition of RNA in comparison to DNA.

RNA plays a crucial role as a catalyst inside the ribosome and mediates many transactions in the cell.<sup>7</sup> In this context, RNA has transformed from a molecule with a minor role in protein synthesis to an important player in molecular biology.<sup>8,9</sup> Thus, the development of RNA detection and recognition technology is gaining immense importance for having an enormous impact on molecular biology and medicine.<sup>10</sup> The direct visualization of nucleic acids *in vivo* can provide information about the location, kinetics and function of these biomolecules, playing a major role in understanding different inter- and intracellular processes.<sup>11</sup> Moreover, dynamic quantitative detection of RNA is a vital subject in neurotoxin and cancer biology as variation in RNA abundance is related to gene expression.<sup>12</sup>

Given the diversity of RNA functions, small fluorescent probes that selectively bind to RNA would be a highly efficient approach for therapeutic intervention. Small cationic imaging probes are frequently applied in biological research.<sup>13-17</sup> However, the problem of these small cationic fluorescent probes is that they generally have better affinity for DNA over RNA.<sup>13-15</sup> Chang and co-workers reported two small molecules for RNA detection; however, relatively small differences in affinity were observed between RNA and DNA.<sup>13-15</sup> Yoon and coworkers reported a pyrene based neutral probe which is selective towards RNA compared to DNA but is unselective towards other nucleotides present in biological fluids.<sup>16</sup> Shirinifar *et al.* reported small naphthalene-imidazolium based cationic cyclophane which can selectively detect RNA over DNA in living cells.<sup>17</sup> However, the recognition mechanism of the reported probes for RNA was undefined. In this regard, a concise and explicit binding mechanism for fluorescence

<sup>a</sup>Pohang University of Science and Technology, Pohang 790-784, Korea<sup>b</sup>Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Korea.  
E-mail: kimks@unist.ac.kr

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5sc03740a

‡ Authors contribute equally in this work.



sensing of RNA is essential for further development of fluorescent molecules showing high selectivity and specificity towards RNA.

For the sake of the present study we synthesized naphthalene-, anthracene- and pyrene-based probes (Fig. 1).  $^1\text{H}$  NMR, fluorescence titration, and circular dichroism (CD) experiments were performed to explain the binding stoichiometry and structures. Molecular dynamics (MD) simulations were performed to confirm whether the RNA structure violates the neighbor-exclusion principle within several nanoseconds, which is the time scale for fluorescence. Additionally, density functional theory (DFT) and time-dependent DFT (TD-DFT) methods were used to elucidate the binding and fluorescence mechanisms using stacking nucleobases which exist in the tRNA of baker's yeast: A-A, A-C, A-G, A-U, C-G, C-U, G-G, G-U and U-U pairs (A: adenine; C: cytosine; G: guanine; U: uracil; Fig. S28†). Additional experiments using a probe that replaced the imidazolium moiety with a triethyl amino group clarify the role of charged groups along with the structures suggested by DFT results. Consequently, both the experimental evidence and theoretical calculations indicate the violation of the well-known neighbor-exclusion principle: naphthalene-based small cationic hosts fit into every individual stacking nucleobase with  $\pi$ - $\pi$  interactions<sup>18</sup> between the fluorophores and the nucleobases. It is also shown that the entire structure is aided by ionic hydrogen bonding<sup>19-21</sup> between the positively charged moiety and the ribose-phosphate backbone, resulting in selectivity for RNA over DNA. This study on the mode of binding is essential and progressive as no concrete analysis has been reported yet.

## 2. Results and discussion

### 2.1. Fluorescence studies

The details of synthesis of probes 1-5 are described in the ESI (Scheme S1 and S2†).<sup>17,22-25</sup> Fluorescence studies have been conducted for selective recognition of RNA because of its simplicity, quick response, and high sensitivity at low concentrations. The fluorescence results of probes 1-5 were checked in aqueous solution at pH 7.4 (10 mM HEPES buffer). The final concentration of tRNA from baker's yeast and RNA from *torula* yeast was determined spectrometrically ( $\epsilon_{260} = 9250 \text{ M}^{-1} \text{ cm}^{-1}$ , expressed as molarity of phosphate groups).<sup>26</sup> Probes 1-3 display lower fluorescence emissions ( $\lambda_{\text{max}} = 402 \text{ nm}$ ) when irradiated at 350 nm (Fig. 2).<sup>17</sup> Due to the presence of the quenching effect of the imidazolium moieties in probes 1-2,

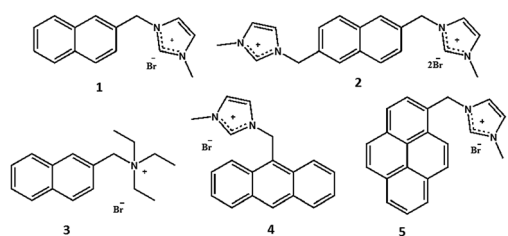


Fig. 1 Probes 1-5 (naphthalene based probes 1-3, anthracene based probe 4 and pyrene based probe 5).

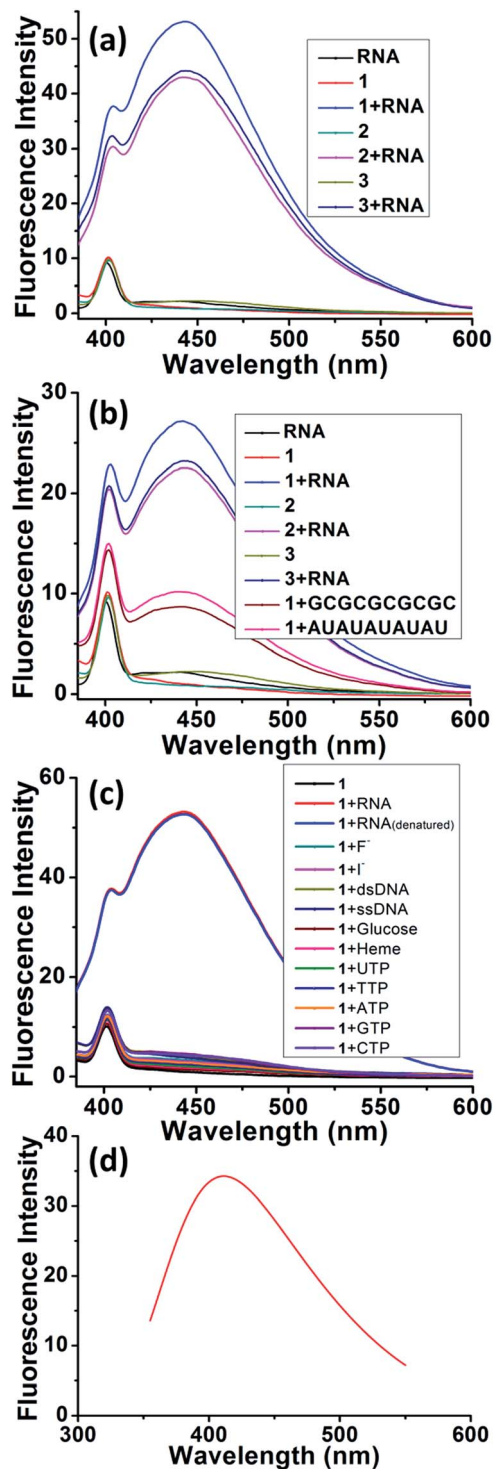


Fig. 2 Experimental and theoretically calculated fluorescence spectra. (a) Fluorescence of tRNA from baker's yeast (10  $\mu\text{M}$ ), probes 1-3 (10  $\mu\text{M}$ ) and probes 1-3 with tRNA from baker's yeast. (HEPES buffer pH = 7.4, slit width = 5 nm) (b) fluorescence of RNA from *torula* yeast (10  $\mu\text{M}$ ), probes 1-3 (10  $\mu\text{M}$ ), probes 1-3 with RNA from *torula* yeast and fluorescence of probe 1 with tRNA (GCGCGCGCGC and AUUAUAUAU). (HEPES buffer pH = 7.4, slit width = 5 nm) (c) fluorescence of probe 1 (10  $\mu\text{M}$ ) in the presence of RNA, denatured RNA,  $\text{F}^-$ ,  $\text{Cl}^-$ , heme, glucose, ssDNA, dsDNA, UTP, TTP, ATP, GTP and CTP (10 equiv.) (HEPES buffer pH = 7.4, slit width = 5 nm). (d) Fluorescence results from TD-DFT calculations.



negligible fluorescence emission is observed (quantum yield = 0.04 and 0.06 for probes 1 and 2, respectively) while a quantum yield of 0.05 is observed in the case of probe 3.<sup>17,22</sup> tRNA from baker's yeast and RNA from *torula* yeast exhibit negligible fluorescence (Fig. 2) but turn-on fluorescence ( $\lambda_{\text{max}} = 443$  nm) is observed in the emission spectrum when the probes were treated with tRNA from baker's yeast (quantum yield = 0.63, 0.59, and 0.52 for probes 1–3, respectively) and RNA from *torula* yeast (quantum yield = 0.41, 0.32 and 0.31 for probes 1–3, respectively). The turn-on fluorescence ( $\lambda_{\text{max}} = 443$  nm) is also observed in the emission spectra for the probe 1 treated with tRNA (GCGCGCGCGC) with a quantum yield = 0.22 and tRNA (AUAUAUAUAU) with a quantum yield = 0.21. This indicates that the fluorescence sensing does not depend on the structural skeleton (Fig. 2). Probe 3 (where the imidazolium group is replaced by a triethyl amino group) gives a similar fluorescence enhancement, indicating that the imidazolium group just gives an electrostatic interaction and has no effect on recognition. Probe 4 exhibits monomer emissions at 398, 421 and 444 nm,<sup>22</sup> while probe 5 shows monomer emissions at 379, 398 and 419 nm when irradiated at 350 nm.<sup>27</sup> Both probes 4 and 5 give decreased fluorescence in the monomer when exposed to tRNA, indicating that the naphthalene moiety is responsible for the fluorescence enhancement and hence for the selective recognition of tRNA (Fig. S11†). The new peak at  $\sim 443$  nm upon binding of probes 1–3 with tRNA from baker's yeast, RNA from *torula* yeast, tRNA (GCGCGCGCGC) and tRNA (AUAUAUAUAU) is attributed to excimer formation.<sup>17</sup> The absorption spectra of probes 1–3 showed a broad peak  $\sim 265$  nm in the presence of the naphthalene moiety, which became sharper with a distinct blue-shifting of  $\sim 6$  nm upon interaction of tRNA with baker's yeast (Fig. S12†). On the other hand, the absorption spectra of probes 4 and 5 upon binding with tRNA from baker's yeast show negligible change (Fig. S13†). Probe 1 shows almost insignificant fluorescence enhancement when exposed to  $\text{F}^-$ ,  $\text{I}^-$ , double-stranded (ds) DNA, single-stranded (ss) DNA, glucose, heme, UTP, TTP, ATP, GTP and CTP (Fig. 2c), indicating the high selectivity of naphthalene-based probe 1 towards recognition of tRNA and denatured tRNA over other anions and nucleotides.

Fluorescence titration experiments of probes 1–3 upon binding with tRNA from baker's yeast and RNA from *torula* yeast were conducted to investigate the binding phenomenon (Fig. S14–S19†). Probes 1–3 display a 1 : 1 binding stoichiometry with tRNA from baker's yeast and RNA from *torula* yeast (the concentration of RNA was determined spectrometrically and expressed as the molarity of phosphate groups),<sup>26</sup> suggesting that the imidazolium/triethyl amino group of each probe molecule binds to each phosphate unit of the RNA backbone, through primary electrostatic interaction. Simultaneously, this allows the naphthalene moiety of each probe molecule to have secondary interactions with the stacking nucleobases of RNA. All these observations are further supported by the fact that excimer formation arises from interstitial  $\pi$ – $\pi$  stacking<sup>22</sup> between naphthalene moiety and each stacking pair of RNA.<sup>6</sup> Only one imidazolium group might be involved in binding as evident from the 1 : 1 stoichiometry (Fig. S15 and S18†) in the case of probe 2. Binding constants<sup>28,29</sup> ( $\sim 10^4 \text{ M}^{-1}$ ) and detection

limits<sup>30</sup> ( $\sim 8 \times 10^{-6} \text{ M}$ ) of probes 1–3 were calculated and are summarized in Table 1. Fluorescence titration results were also subjected to a Scatchard plot to calculate binding constants (Fig. S14–S19† and Table 1).<sup>31</sup> The results are almost comparable to the binding constants calculated based on 1 : 1 binding stoichiometry between the phosphate group of RNA and probes 1–3 (Table 1). Hence, this strengthens our argument that the imidazolium/triethyl amino group of each probe molecule binds to each ribose-phosphate unit of RNA.<sup>31</sup>

In order to confirm the fluorescence result, we carried out DFT and TD-DFT calculations<sup>32</sup> on intercalation model systems in which probe 1 is sandwiched between two nucleobases that are connected by a ribose-phosphate backbone (Table S1†). In the ground state, all the structures have perfect triple stacking with  $\pi$ – $\pi$  interactions. The molecular orbitals (MOs) responsible for the vertical excitation (absorption), mostly the highest occupied MOs (HOMOs) and the lowest unoccupied MOs (LUMOs), are delocalized over the naphthalene and one of the bases. Due to the large spatial overlaps between the HOMOs and LUMOs and the similar structures, no significant difference is observed in wavelength and oscillator strength for absorption between different structures with an intercalated naphthalene moiety. However, the rotation of either one of the two bases or the naphthalene moiety leads to breaking of the spatial overlap between HOMO and LUMO at the 1<sup>st</sup> excited state minimum, which is essential for de-excitation and fluorescence. This is followed by wide-ranging wavelength values for the 1<sup>st</sup> excited state optimized structures; the values span from 393 to 503 nm (Fig. 3 and S22†). The structures can be classified in terms of oscillator strength. Fig. S20 and S22† explain that A–1–A, A–1–C and A–1–G have the delocalized HOMOs over one of the bases and the naphthalene moiety, resulting in significant overlap with LUMOs mostly localized on the naphthalene moiety which showed high oscillator strength. On the other hand, other structures present smaller oscillator strength due to localized HOMOs and LUMOs either on naphthalene or one of the bases, so-called charge-transfer (CT) de-excitation. The exceptions are C–1–G and C–1–U: oscillator strength  $< 0.01$  for the former with delocalized HOMO and localized LUMO, but  $> 0.01$  for the latter even with CT excitation. The data could be merged together to reproduce the fluorescence result. Considering both Doppler broadening for the finite width to the spectral lines and the number of each stacking pair, the theoretical fluorescence matches experiment well (Fig. 2d and ESI†). This strongly indicates that not only specific stacking pairs but all the intercalation structures are involved in the fluorescence ranging 425–450 nm.

## 2.2. Circular dichroism (CD) studies

To obtain insight into the conformational change of the RNA structure and binding mechanism upon interaction, probe 1 was selected for circular dichroism (CD) studies.<sup>33</sup> tRNA from baker's yeast (1 mM) shows positive ellipticity centred at 276.5 nm (Fig. 4a and S23†) while RNA from *torula* yeast shows positive ellipticity centred at 290.5 nm (Fig. S23†) (due to the stacking interactions between the stacking pairs and the helical



Table 1 Results of binding stoichiometries, binding constants and detection limits of probes 1–3 with RNA

Probe	Binding stoichiometry	Binding constant ( $M^{-1}$ )	Binding constant ( $M^{-1}$ ) (Scatchard plot)	Detection limit (M)
<b>tRNA from baker's yeast</b>				
1	1 : 1	$1.26 \pm 0.1 \times 10^4$	$1.25 \pm 0.03 \times 10^4$	$7.76 \times 10^{-6}$
2	1 : 1	$1.15 \pm 0.1 \times 10^4$	$1.15 \pm 0.02 \times 10^4$	$7.72 \times 10^{-6}$
3	1 : 1	$1.12 \pm 0.09 \times 10^4$	$1.14 \pm 0.03 \times 10^4$	$7.51 \times 10^{-6}$
<b>tRNA from torula yeast</b>				
1	1 : 1	$1.03 \pm 0.1 \times 10^4$	$1.04 \pm 0.02 \times 10^4$	$8.10 \times 10^{-6}$
2	1 : 1	$1.07 \pm 0.09 \times 10^4$	$1.02 \pm 0.03 \times 10^4$	$7.86 \times 10^{-6}$
3	1 : 1	$1.01 \pm 0.08 \times 10^4$	$1.05 \pm 0.02 \times 10^4$	$7.98 \times 10^{-6}$
<b>CD results (tRNA from baker's yeast)</b>				
1	1 : 1	$0.98 \pm 0.08 \times 10^4$	$0.99 \pm 0.02 \times 10^4$	$8.45 \times 10^{-6}$
<b>CD results (tRNA from torula yeast)</b>				
1	1 : 1	$0.97 \pm 0.08 \times 10^4$	$1.00 \pm 0.02 \times 10^4$	$8.66 \times 10^{-6}$

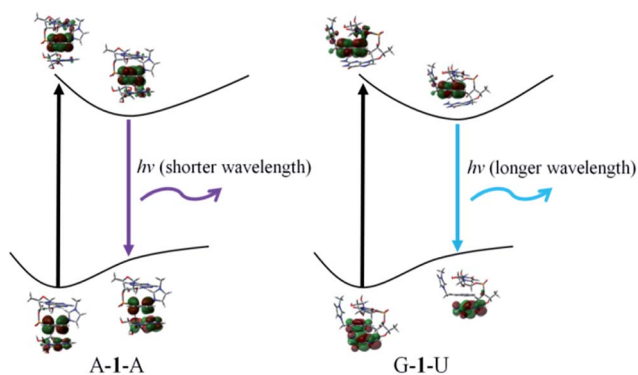


Fig. 3 Schematic description of the fluorescence mechanism.

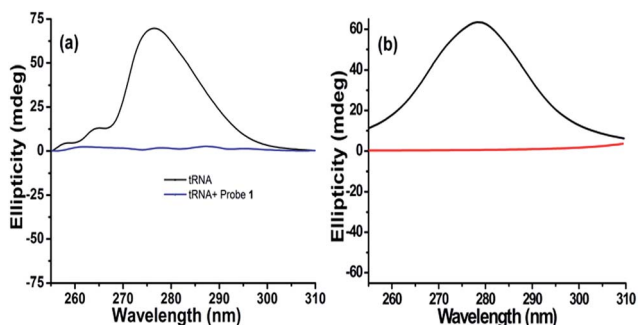


Fig. 4 Experimental and theoretically calculated CD data. (a) CD of tRNA from baker's yeast (2 mM) with and without probe 1 (4.5 equiv.). HEPES buffer pH = 7.4. (b) Computed CD results of the tRNA fragment with 10 nucleotides: helical structure without probe 1 (black) and stretched structure with probe 1 (red).

structure that provide an asymmetric environment for the bases). Addition of probe 1 into tRNA from baker's yeast and RNA from *torula* yeast solution results in a decrease in ellipticity until it becomes almost zero (Fig. 4a and S23<sup>†</sup>).<sup>34</sup> CD titration of tRNA from baker's yeast and RNA from *torula* yeast shows 1 : 1

stoichiometry and the binding constant is calculated assuming 1 : 1 binding stoichiometry (Fig. S24 and S25<sup>†</sup>). CD titration results are also subjected to Scatchard plot to calculate the binding constants (Fig. S24 and S25<sup>†</sup>) and the outcomes are comparable to those calculated based on 1 : 1 binding stoichiometry between the phosphate group of RNA and probes 1–3 (Table 1). Thus, this strengthens our argument that the imidazolium/triethyl amino group of each probe molecule binds to each ribose-phosphate unit of RNA.<sup>31</sup> Based on these observations, we propose that interaction of probe 1 causes tRNA from baker's yeast and RNA from *torula* yeast to unfold their secondary structures, exploiting strong binding between RNA and probe 1.<sup>35,36</sup> The phenomenon has been supported by fluorescence studies where tRNA and denatured tRNA with probe 1 give almost the same response (Fig. 2c). Fluorescence (Fig. S14–S19e and f<sup>†</sup>) and circular dichroism (Fig. S24 and S25e and f<sup>†</sup>) results were also subjected to neighbor exclusion model (proposed by Schellman and Reese) and a cut off of experimental points was observed well below  $\theta = 0.5$  indicating that the neighbour exclusion principle is upheld in RNA recognition by probes 1–3.<sup>37</sup>

The computed CD result for the model, shown in Fig. 4b, further confirms our speculation that the stretched RNA structure is responsible for the flat signal in CD experiments. RNA used here is only a local part from the entire tRNA of baker's yeast with a limited number of nucleotides, showing that the wavelength at the maximum ellipticity ( $\sim 278$  nm) is very close to the experimental result (Fig. 4b). Moreover, the signal for the stretched RNA structure with probe 1 is almost flat as in Fig. 4a. Therefore, we are convinced that the CD results can be attributed to the involvement of probe molecules in the stretch of RNA, where probe molecules are intercalated into every base stacking site, violating the neighbor-exclusion principle. This argument is consistent with the basic insight of hydrophobicity: all the hydrophobic naphthalene fluorophores try to avoid exposure to hydrophilic environment due to lack of attraction with hydrophiles. There are only two possibilities for naphthalene to escape the hydrophilic environment, either aggregation



of the probe molecules or intercalation in between nucleobases both using  $\pi$ - $\pi$  interaction. However in this case, the charged moiety (imidazolium or triethyl amino group) prevents the former because of its solubility in water.

### 2.3. $^1\text{H}$ NMR studies

Probe **1** was selected in order to monitor the physical interaction through  $^1\text{H}$  NMR experiment. tRNA (GCGCGCGCGC) and tRNA (AUAUAUAUAU) were used to investigate which nucleobase is responsible for the interaction with the naphthalene moiety of probe **1** (see ESI† for details). A 2D NOESY experiment of probe **1** with tRNA from baker's yeast was also recorded in order to investigate the proposed binding pattern. The relatively weak NOE correlation between the naphthalene moiety of probe **1** and the nucleobases of RNA reveals that the naphthalene moiety is in close vicinity to the nucleobases of RNA (Fig. S26†). Downfield shifts associated with splitting of naphthalene protons and upfield shifts of RNA protons (Fig. S27 and S28†) suggest that each nucleobase of RNA is involved in  $\pi$ - $\pi$  stacking interactions with the naphthalene moiety of probe **1** causing excimer formation and fluorescence with a broad peak centred around 425–450 nm. This was further strengthened by 1 : 1 binding stoichiometry of probe **1** with the phosphate groups of RNA. Fluorescence, circular dichroism and  $^1\text{H}$  NMR data demonstrate that every stacking pair is involved in intercalation and fluorescence, in contrast to the previous conjecture that one naphthalene based cyclophane binds to one RNA molecule.<sup>17</sup>

### 2.4. Molecular dynamics (MD) simulations

Furthermore, our MD simulation results support our proposal about unfolding of RNA due to intercalation of probes. We performed MD simulations for RNA with 10 nucleotides in which the naphthalene moiety in probe **1** is located at every intercalation site. The majority of naphthalene moieties in the structure, where some imidazolium moieties interact with phosphate and the others interact with ribose's 2'-OH, maintain the initial intercalation form for 15 ns (Fig. 5). On the other hand, if imidazolium moieties interact only with either the phosphate or the 2'-OH of ribose, the structures with intercalating probe molecules become highly unstable, lasting at most for several nanoseconds (Fig. S30 and S31†). Likewise, the MD simulation of the DNA fragment with the same sequence in which intercalated probe molecules interact only with the phosphate backbone shows the dissolution of almost all the intercalation structures within 0.4 ns (Fig. S32†). This is clearly shown from the root-mean-squared distance and deviation data between a probe molecule and stacking nucleobases for RNA-probe and DNA-probe systems (Fig. S33 and S34†). While the probe molecule in the RNA system maintains the intercalation structure even after 20 ns, the one in the DNA system is solvated out within sub-nanoseconds. These results emphasize the difference between DNA and RNA. The lack of 2'-OH in DNA forces imidazolium moieties to interact only with phosphate backbones for at most a few nanoseconds, making the whole intercalation structure unstable. On the other hand, RNA provides two different options for interaction, 2'-OH and the

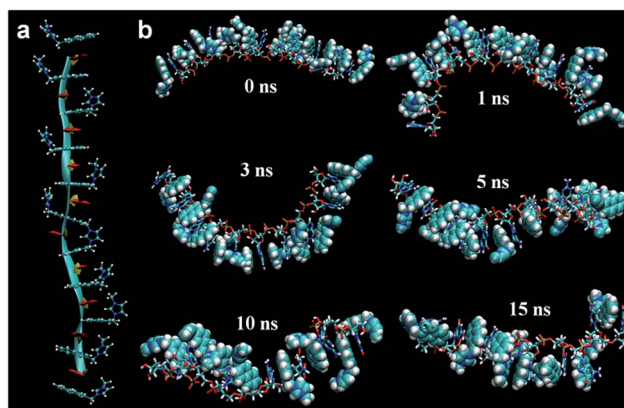


Fig. 5 MD simulation results (a) schematic description of tRNA fragment of 10 nucleotides interacting with probe **1**: naphthalene moieties at intercalation sites and imidazolium moieties interacting with phosphate backbone and 2'-hydroxyl group (2'-OH) of ribose (ribbon: phosphate backbone; ball-and-stick: probe **1**; yellow: ribose; red: nucleobase). (b) Snapshots of 15 ns MD simulation in NPT ensemble of the corresponding model (stick: RNA; vdW: probe **1**). Water molecules are removed for clarity.

phosphate backbone, directing imidazolium moieties toward the opposite direction for hydrogen bonding. This offers additional stability for intercalation structures, lasting several tens of nanoseconds. Therefore, we can conclude that once the probe molecules are intercalated, they stay where they are in the initial form for several tens of nanoseconds corresponding to the time scale of fluorescence,  $\sim 1$ –100 ns. Along with the NMR experiment data, this indicates that the intercalation structures are responsible for the fluorescence results.

## 3. Conclusion

We have shown that the neighbor-exclusion principle is violated in RNA by naphthalene based cationic probes. The control experiments demonstrate that only the naphthalene moiety is small enough to be inserted into intercalation sites of RNA, unlike pyrene and anthracene. Furthermore, fluorescence titration, CD and  $^1\text{H}$  NMR experiments infer that not only specific bases but all the stacking pairs are responsible for the fluorescence, having  $\pi$ - $\pi$  interactions with the probes. The MD simulation results reveal stable intercalation structures where imidazolium moieties interact with the two H-bonding acceptors (the negatively charged oxygen of phosphate and the 2'-OH of ribose) both present in RNA, providing selectivity towards RNA for the cationic moieties of the probes as compared to DNA which only has the phosphate backbone. The computational results support the CD experimental results, suggesting that RNA is stretched by the intercalation of probe molecules. Additionally, we computationally reproduced fluorescence, providing the fluorescence mechanism and supporting the involvement of all intercalation structures in the fluorescence. The breaking of the spatial overlap between HOMO and LUMO at each 1<sup>st</sup> excited state minimum gives charge transfer driven de-excitations corresponding to fluorescence at 425–450 nm.



Overall, we have proposed an effective strategy for RNA recognition, a small fluorophore for facile intercalation with at least one cationic moiety for hydrogen bonding, which is straightforward for both synthesis and further analysis.

## 4. Experimental section

### 4.1. Materials and methods

The synthesis of compounds 1–5 is described in the ESI.† Synthesized compounds (1–5) were fully characterized with standard spectroscopic techniques. Imidazole, 2,6-bis-(bromomethyl) naphthalene, 2-bromomethyl naphthalene, 9-bromomethyl anthracene, 1-bromomethyl pyrene and triethyl amine were purchased from Aldrich and were used as such. Sodium salts of ATP, GTP, CTP, TTP and UTP, heme, glucose, dsDNA (from Calf Thymus), RNA (from baker's yeast) and RNA (from *torula* yeast) were also purchased from Aldrich and used without further purification. Tetrabutyl ammonium salts of F<sup>-</sup> and I<sup>-</sup> were also purchased from Aldrich and used without further purification. tRNA (GCGCGCGCGC) and tRNA (AUAUAUAUAU) were purchased from XIDT and used as such.

**Fluorometric analysis.** Stock solutions of compounds 1, 2, 3, 4, and 5 (1 mM) were prepared at pH 7.4 in 0.01 M HEPES buffer water mixture and used in the preparation of titration solutions by appropriate dilution up to 10 μM. Aliquots of ATP, GTP, CTP, UTP, TTP, heme, glucose, dsDNA, ssDNA, RNA from baker's yeast, RNA from *torula* yeast, tRNA (GCGCGCGCGC) and tRNA (AUAUAUAUAU) in 0.01 M HEPES buffer water mixture were then injected into the sample solution through a rubber septum in the cap. dsDNA solution was heated to 90 °C and rapidly cooled at 4 °C in order to denature it.<sup>26</sup> Similarly tRNA from baker's yeast was also denatured by heating its solution to 90 °C and rapidly cooling at 4 °C.<sup>26</sup>

**Circular dichroism studies.** The CD spectra were collected at room temperature (25 ± 0.2 °C) using a Jasco made J-815 CD spectropolarimeter. Scans were from 350 to 200 nm with a resolution of 1 nm, with data sampling every 5 s. The 1 cm cell contained 1 mM solution of the RNA in 0.01 M HEPES buffer (pH 7.4).<sup>26</sup> 0.1 M solution of probe 1 was also prepared in 0.01 M HEPES buffer (pH 7.4). CD spectra were then recorded with pure RNA and with addition of specific amount of probe 1.

**MD simulations.** The atomic coordinates of tRNA of baker's yeast were taken from the Protein Data Bank (PDB) (entry 3EPK: tRNA of baker's yeast entangled with eukaryotic dimethylallyl-transferase; Fig. S29a†).<sup>32</sup> Missing hydrogen atoms were added using the psfgen module implemented in NAMD program.<sup>38</sup> 10 nucleotides were obtained from this structure (sequenced as AGACGACGCG) and their backbones were stretched for probe molecules to be intercalated in between bases. Ribose groups were patched up by deoxyribose groups for DNA construction. The topology and parameters of probe 1 were constructed using the CHARMM general force field (CGenFF) program after geometry optimization at the M06/6-31G\* level.<sup>39–41</sup> The structures are given in Fig. 5 and S30–S32.† The entire structure consisting of the RNA/DNA fragment and probe 1 molecules was then soaked into TIP3P water box. After minimizing the box for 10 heating from 0 to 295 K for 10 ps, we equilibrated the

structure in NPT ensemble for 1 ns using the Nose-Hoover Langevin piston pressure control.<sup>42</sup> Next, we performed MD simulations using the NAMD program and CHARMM36 force field<sup>43</sup> with periodic boundary conditions and particle-mesh Ewald (PME) full electrostatics.<sup>44</sup> Coarse PME grid was used to speed up the simulations. By applying Langevin forces the temperature was maintained at 295 K. Van der Waals energies were calculated using cutoff of 12 Å. The MD simulations were performed for 15 ns.

**(TD-)DFT computations.** The single point TD-DFT computations were performed for a structure with probe 1 at the intercalation sites obtained from the previous MD simulation and an original helical structure (from the original PDB file) for theoretical CD. The M06/6-31G\* level of theory was employed with conductor polarizable continuum model (CPCM).<sup>45,46</sup>

For the in-depth study of the fluorescence mechanism, we performed detailed (TD-)DFT computations further. We acquired some representative base-base stacking fragments including ribose-phosphate backbone from the above mentioned PDB file; we could obtain A–A, A–C, A–G, A–U, C–G, C–U, G–G, G–U and U–U stacking pairs. Then, probe 1 was intercalated in between each stacking base due to its smallest size among the probes we synthesized. To deal with the systems more realistically, some water molecules were added into the first solvation shell around bases along with the CPCM. We optimized the ground and 1<sup>st</sup> excited states and measured the energies and oscillator strengths. For CD calculations, we used M06/6-31G\* level of theory for (TD-)DFT computations since it gives a similar UV-visible absorption spectrum (maximum oscillator strength of 0.91 at wavelength of 215.5 nm) of the probe 1 to the experiment (maximum intensity at wavelength of 224.3 nm).<sup>47–50</sup> All calculations were performed using Gaussian 09 program.<sup>51</sup>

In order to reproduce the fluorescence results, practically it is too time-consuming to compute all the pairs explicitly using the method explained above. Instead, we made approximations; the wavelengths and oscillator strengths of the same stacking nucleobases are the same. Also, one more important point is the Doppler broadening, which brings out a finite width to the spectral lines.<sup>52</sup> After considering the Doppler broadening, the oscillator strength values obtained from a certain stacking pair were multiplied by the number of the pair in the given RNA structure, which comprises 1 A–A, 6 A–C, 11 A–G, 4 A–U, no C–C, 9 C–G, 10 C–U, 4 G–G, 9 G–U and 3 U–U stacking pairs.

## Acknowledgements

This work was supported by the NRF (National honor scientist program: 2010-0020414) and KISTI (KSC-2014-C3-019).

## Notes and references

- 1 D. M. Crothers, *Biopolymers*, 1968, **6**, 575–584.
- 2 A. A. Greschner, K. E. Bujold and H. F. Sleiman, *J. Am. Chem. Soc.*, 2013, **135**, 11283–11288.
- 3 E. D. Horowitz, S. Lilavivat, B. W. Holladay, M. W. Germann and N. V. Hud, *J. Am. Chem. Soc.*, 2009, **131**, 5831–5838.



- 4 L. D. Williams, M. Egli, Q. Gao and A. Rich, DNA intercalation: helix unwinding and neighbor-exclusion, *Structure & Function: Nucleic Acids*, ed. R. H. Sarma and M. H. Sarma, Adenine Press, 1992, vol. 1.
- 5 P. A. Kollman and K. M. Merz Jr, *Acc. Chem. Res.*, 1990, **23**, 246–252.
- 6 S. N. Rao and P. A. Kollman, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 5735–5739.
- 7 G. F. Joyce, *Nature*, 2002, **418**, 214–221.
- 8 I. Tinoco Jr and C. Bustamante, *J. Mol. Biol.*, 1999, **293**, 271–281.
- 9 G. North, *Nature*, 1987, **328**, 18–19.
- 10 S. Valastyan, F. Reinhardt, N. Benaich, D. Calogrias, A. M. Szasz, Z. C. Wang, J. E. Brock, A. L. Richardson and R. A. Weinberg, *Cell*, 2009, **137**, 1032–1046.
- 11 A. P. Silverman and E. T. Kool, *Trends Biotechnol.*, 2005, **23**, 225–230.
- 12 M. Lucanic, J. Graham, G. Scott, D. Bhaumik, C. C. Benz, A. Hubbard, G. J. Lithgow and S. Melov, *Aging*, 2013, **5**, 394–411.
- 13 Q. Li, Y. K. Kim, J. Namm, A. Kulkarni, G. Rosania, Y. H. Ahn and Y. T. Chang, *Chem. Biol.*, 2006, **13**, 615–623.
- 14 C. J. O'Connor, H. S. G. Beckmann and D. R. Spring, *Chem. Soc. Rev.*, 2012, **41**, 4444–4456.
- 15 S. Feng, Y. K. Kim, S. Yang and Y. T. Chang, *Chem. Commun.*, 2010, **46**, 436–438.
- 16 Y. Liu, E. J. Jun, G. Kim, A.-R. Lee, J.-H. Lee and J. Yoon, *Chem. Commun.*, 2014, **50**, 2505–2507.
- 17 B. Shirinfar, N. Ahmed, Y. S. Park, G.-S. Cho, I. S. Youn, J.-K. Han, H. G. Nam and K. S. Kim, *J. Am. Chem. Soc.*, 2013, **135**, 90–93.
- 18 (a) K. S. Kim, P. Tarakeshwar and J. Y. Lee, *Chem. Rev.*, 2000, **100**, 4145–4185; (b) C. A. Hunter and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1990, **112**, 5525–5534; (c) C. A. Hunter, K. R. Lawson, J. Perkins and C. J. Urch, *J. Chem. Soc., Perkin Trans. 2*, 2001, **5**, 651–669; (d) R. S. Paton and J. M. Goodman, *J. Chem. Inf. Model.*, 2009, **49**, 944–955; (e) L. Rasheed, M. Yousuf, I. S. Youn, T. Yoon, K. Y. Kim, Y. K. Seo, G. Shi, M. Saleh, J. H. Hur and K. S. Kim, *Chem.–Eur. J.*, 2015, **21**, 16349–16353.
- 19 J. Černý, M. Kabeláč and P. Hobza, *J. Am. Chem. Soc.*, 2008, **130**, 16055–16059.
- 20 H. Ihm, S. Yun, H. G. Kim, J. K. Kim and K. S. Kim, *Org. Lett.*, 2000, **4**, 2897–2900.
- 21 J. Yoon, S. K. Kim, N. J. Singh and K. S. Kim, *Chem. Soc. Rev.*, 2006, **35**, 355–360.
- 22 M. Yousuf, N. Ahmed, B. Shirinfar, V. M. Miriyala, I. S. Youn and K. S. Kim, *Org. Lett.*, 2014, **16**, 2150–2153.
- 23 F.-J. Chen, Y. Xu and H.-B. Du, *Angew. Chem., Int. Ed.*, 2014, **53**, 9592–9596.
- 24 D. Jiao, F. Biedermann and O. A. Scherman, *Org. Lett.*, 2011, **13**, 3044–3047.
- 25 J. J. Henkelis, C. J. Carruthers, S. E. Chambers, R. Clowes, A. I. Cooper, J. Fisher and M. J. Hardie, *J. Am. Chem. Soc.*, 2014, **136**, 14393–14396.
- 26 R. Mart, C. N. N'soukpoé-Kossi, D. M. Charbonneau, L. Kreplak and H.-A. Tajmir-Riahi, *Nucleic Acids Res.*, 2009, **37**, 5197–5207.
- 27 Z. Xu, N. J. Singh, J. Lim, J. Pan, H. N. Kim, S. Park and K. S. Kim, *J. Am. Chem. Soc.*, 2009, **131**, 15528–15533.
- 28 (a) K. A. Connors, *Binding Constants: The Measurement of Molecular Complex Stability*, Wiley, New York, 1987; (b) N. Ahmed, B. Shirinfar, I. S. Youn, M. Yousuf and K. S. Kim, *Org. Biomol. Chem.*, 2013, **11**, 6407–6413.
- 29 *OriginLab 8.0*, OriginLab Corporation, Northampton, MA, 2003.
- 30 M. Shortreed, R. Kopelman, M. Kuhn and B. Hoyland, *Anal. Chem.*, 1996, **68**, 1414–1418.
- 31 (a) J. D. McGhee and P. H. von Hippel, *J. Mol. Biol.*, 1974, **86**, 469–489; (b) M. S. Rocha, *Biopolymers*, 2009, **93**, 1–7.
- 32 C. Zhou and R. H. Huang, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 16142–16147.
- 33 Y.-M. Chang, C. K.-M. Chen and M.-H. Hou, *Int. J. Mol. Sci.*, 2012, **13**, 3394–3413.
- 34 A. Das and G. S. Kumar, *Mol. BioSyst.*, 2012, **8**, 1958–1969.
- 35 M. K. Sharon and C. P. Nicholas, *Biochim. Biophys. Acta*, 1997, **1338**, 161–185.
- 36 M. K. Sharon and C. P. Nicholas, *Curr. Protein Pept. Sci.*, 2000, **1**, 349–384.
- 37 J. A. Schellman and H. A. Reese, *Biopolymers*, 1995, **39**, 161–171.
- 38 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
- 39 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. MacKerell Jr, *J. Comput. Chem.*, 2010, **31**, 671–690.
- 40 W. Yu, X. He, K. Vanommeslaeghe and A. D. MacKerell Jr, *J. Comput. Chem.*, 2012, **33**, 2451–2468.
- 41 K. Vanommeslaeghe and A. D. MacKerell Jr, *J. Chem. Inf. Model.*, 2012, **52**, 3144–3154.
- 42 G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.*, 1994, **101**, 4177–4189.
- 43 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell Jr, *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.
- 44 P. F. Batcho, D. A. Case and T. Schlick, *J. Chem. Phys.*, 2001, **115**, 4003–4018.
- 45 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
- 46 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.
- 47 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
- 48 P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta*, 1973, **28**, 213–222.
- 49 M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654–3665.
- 50 V. Rassolov, J. A. Pople, M. Ratner and T. L. Windus, *J. Chem. Phys.*, 1998, **109**, 1223–1229.



- 51 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, revision B.01*, Gaussian, Inc., Wallingford CT, 2009.
- 52 H. Haken and H. C. Wolf, *The Physics of Atoms and Quanta: Introduction to Experiments and Theory*, Springer-Verlag, Berlin, 5th edn, 1996.

