Modeling of Canonical and C2'-O-thiophenylmethyl Modified hexamers of RNA. Insights into the Nature of Structural Changes and Thermal Stability.

Journal:	*New Journal of Chemistry*
Manuscript ID	NJ-ART-04-2018-001739
Article Type:	Paper
Date Submitted by the Author:	10-Apr-2018
Complete List of Authors:	Dzowo, Yannick; University of Colorado Denver, Chemistry Wolfbrandt, Carly; University of Colorado Denver, Chemistry Resendiz, Marino; University of Colorado Denver, Chemistry WANG, Haobin; University of Colorado Denver, Chemistry;

Modeling of Canonical and C2'-O-thiophenylmethyl Modified hexamers of RNA. Insights into the Nature of Structural Changes and Thermal Stability.

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Yannick Kokouvi Dzowo, Carly Wolfbrandt, Marino J. E. Resendiz* and Haobin Wang*

Modifying the ribose ring or the nucleobase within oligonucleotides of RNA is a strategy that has regained attention due to its potential to provide chemical stability, increase affinity towards a target, or introduce groups with bioorthogonal reactivity, among other applications. However introducing groups via synthetic routes can be a daunting process and may result in unexpected physical properties and structural variations. In addition, characterizing the effects of a modification within an oligonucleotide can be time consuming and expensive, e.g., NMR, crystallography. Thus the development of models that are accurate, reasonably priced and carried out in short periods of time is essential to the development of this field. Herein we report on the development and validation of computational models that explain the nature of the structural and stability changes induced by the presence of a C2'-O-methylthiophene group at various positions within duplexes conformed by two 6-mers of modified RNA as well as random coils containing the same modification(s). Through analysis of the obtained models it was concluded that the presence of the methylthiophene moiety on two strands of RNA induces a conformational change around torsion angle (κ°) that leads to lower values in the pseudorotational angle (δ°), thus resulting in an A-form α-helix with shorter distances between phosphates than those in canonical RNA. These results provide a reasonable explanation for the experimental observations, which display thermal stabilization of the duplexes, and are guiding efforts from our groups to control the structure and function of RNA, i.e., via functionalization of RNA structure to attain thermal stability while retaining structural properties of an A-form duplex.

Introduction

Modified oligonucleotides have been of interest due to their wide potential in various applications, ranging from their use for therapeutic purposes,[1] as biological tools,[2] or as biomaterials,[3] among many others.[4] The structural impact that various modifications have on oligonucleotides of RNA or DNA has been established. However, it often remains a challenge to predict the effects that a new modification will have on structure and function of the target biopolymers. Typical techniques that are used to characterize the structural and functional changes arising from independently generated modifications include X-ray crystallography or NMR and can be carried out at the monomer and the oligomer level. [5-8] Nevertheless, these procedures can be lengthy, unpredictable, or price forbidding. In this respect, the application of computational tools for molecular modeling is an attractive approach given recent developments in theory and computer technology. There have been extensive efforts where computational chemistry has advanced our knowledge and has aided in understanding various processes, from simulations that provide information about local changes[9] to studies aimed at predicting the physical properties of modified oligonucleotides,[10] among some examples. In addition various approaches use electronic structure theory methods, in particular the Kohn-Sham density functional theory (DFT),[11] and have been proven a powerful tool for modeling small to mid-size organic molecules. The insight provided in such studies can be effectively combined with experimental findings to offer a more complete physical picture for the overall process and to guide further experimental investigations.

Along this line the goal of our research is to develop methodology that enables the fast and, sometimes less expensive, assessment in the design of artificial oligonucleotides.[12] Recently, we reported on the duplex stabilization imparted by 2'-*O*-thiophenylmethyl groups in the case that both strands in the duplex are modified at one or

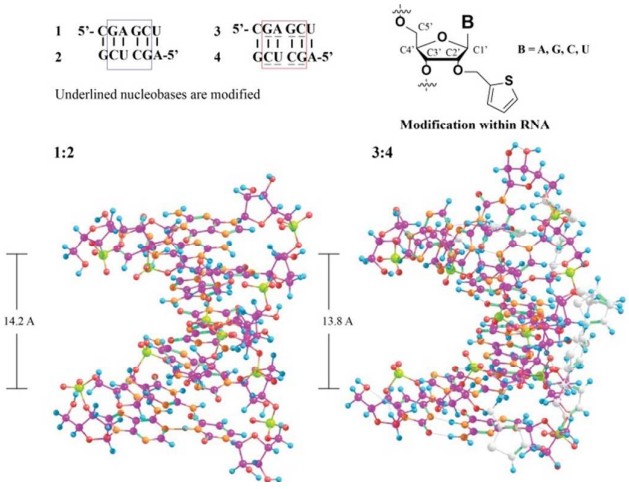

Figure 1. Sequence of the strands modeled in this work (underlined nucleobases represent sites that are modified) along with the structure of the modification, 2-thiophenylmethyl groups (top). Side-by-side comparison of both duplexes, with thiophene units colored in gray for clarity, obtained at the B3LYP-GD3/6-31G* level of theory (bottom). Shown values represent the distance between the furthermost phosphates on both 5'- and 3'-ends of the same strand.

[a] *Department of Chemistry, University of Colorado Denver. 1151 Arapahoe St. Science Building, Denver, CO 80204.*
* To whom correspondence should be addressed.
Electronic Supplementary Information (ESI) available: Plots, formulas, and figures corresponding to all of the measured angles and distances; XYZ coordinates for all structures and calculations; and figures displaying a picture for each calculations are included. See DOI: 10.1039/x0xx00000x

more positions.[13] Complementary to the experimental work, we performed DFT calculations with B3LYP functional and a relatively small basis set (3-21G*) to find the stable geometry for the original and modified duplexes. Analysis of the structural parameters provided the first molecular-level understanding of the duplex stabilization mechanism.

In this work we report a more thorough DFT study on several model molecular systems including the original and modified duplexes as well as the corresponding single strands. We validated our results by taking into account known experimental findings and carried out a careful calibration of the computational results. The latter is done by comparing DFT calculations with different density functionals and basis sets. Using this methodology we assessed the effects of incorporating functional modifications on both strands and further compare to our previous report. The aim is to establish a relatively robust methodology that predicts the effects of various modifications as a function of position, number, and structural nature. Even though our computation is approximate in nature, we found that the major findings are reproducible and that they are in agreement with the experimental data. We further note that molecular mechanics methods that employ empirical force fields, though successful in some other contexts, will not be able to capture the influence of functional modifications on the π–π stacking and other interactions that are quantum mechanical in nature. Therefore, this work contributes to a better microscopic understanding of such molecular interactions and explores the potential stabilization arising from conformational changes at a local (in the vicinity of the modifications) and global (overall structure) level as well as from the proximity between the thiophene rings.

Computational Methods

Density function theory (DFT) calculations were performed using the quantum chemical program Gaussian 09. [14] We employed two functionals, the popular B3LYP hybrid functional and the B97D3 functional which is the B97 functional with Grimme's D3BJ dispersion correction.[15] For the B3LYP functional we also added a variant of including Grimme's dispersion with the original D3 damping function. [16] We have used two basis sets in our investigation, the 3-21G* basis set and the 6-31G* basis set. Due to the large number of atoms, calculations using the latter basis set took considerable computational resources. An implicit solvent approach, the polarizable continuum model (PCM) approach,[17-19] was used to model the solution environment (water in this study). Unrestricted geometry optimizations were carried out for all the models with the PCM approach, which gave much more realistic, compact structures than those relaxed without including solvent effect. In principle there are three functionals (B3LYP, B3LYP-GD3, and B97D3) and two basis sets (3-21G* and 6-31G*), and thus six combinations. In the discussions below we mainly discuss results with the empirical dispersion corrections (B3LYP-GD3 and B97D3), which presumably provide a better description for our molecular systems.

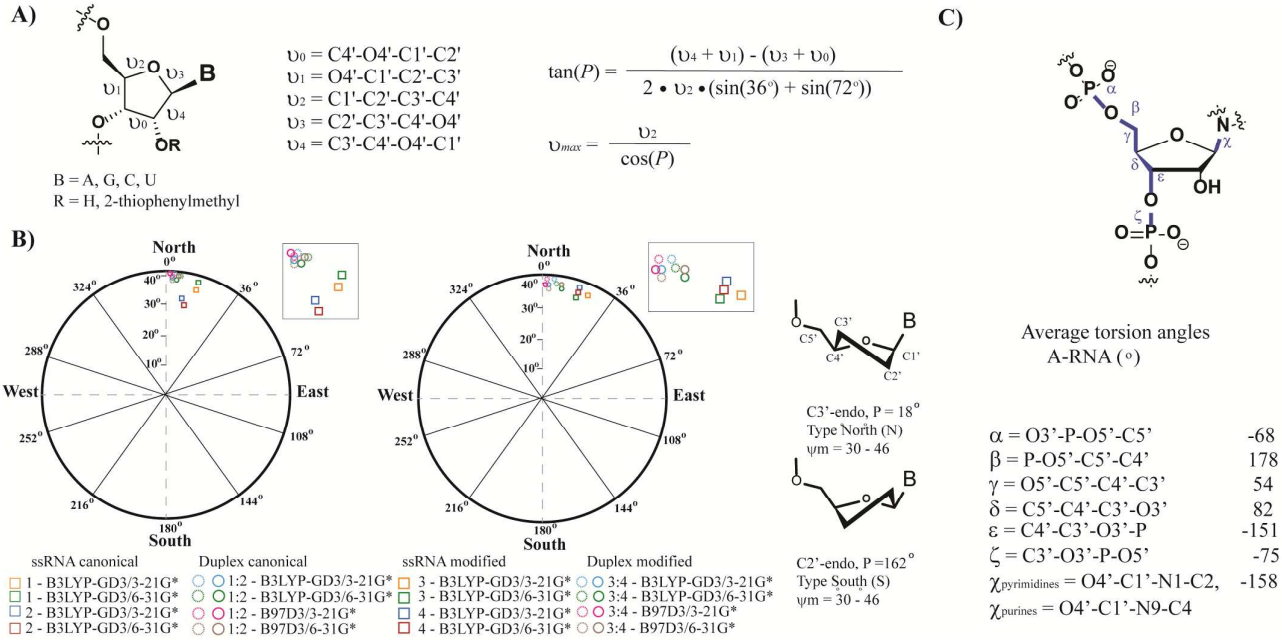

Figure 2. A) Extent of sugar puckering angles and formula; and B) Pseudorotational wheel indicating C3'-endo conformation for both canonical and modified structures in their single and double stranded forms. Contents inside gray boxes highlight the region indicating that all ribose rings are within the conformation expected for a C3'-endo ribose. Squares correspond to single stranded RNA and circles correspond to duplexes (**1** or **3** = dotted lines and **2** or **4** – solid lines). A plot containing the other levels of theory used in this work is provided in the supporting materials. C) Reported values for all of the dihedral angles that provide the necessary structural parameters in oligonucleotides along with the definition for each angle.

We started by building models for which each ribose was in the C3'-endo conformation, canonical RNA pucker. We reasoned that this would be an accurate starting point due to experimental results displaying A-form duplexes with/without modifications. In addition, we carried out calculations that started from various 2'-hydroxyl group orientations, which initially induced structural instability in some cases. However, following optimization, all structures displayed geometries that were consistent with A-form duplexes. The emphasis of the analyses was placed on the double stranded samples due to our experimental data that indicates stabilized structures upon modification of the strands (**3**:**4** displays higher thermal denaturation transitions than **1**:**2**, ref. 13).

Structural analyses of the four model strands **1**-**4** were also carried out on their single stranded (coils) form, with results that displayed larger variations than those observed in the more rigid duplexes. It is worth noting that in order to obtain a more accurate picture of these strands, and further validation of the methodology described herein, a larger sampling set that include various possible conformations is warranted. That is, build models that start from various conformational states followed by the corresponding energy minimization and analysis. This will also allow for a direct comparison obtained from other models that use molecular dynamics.

Data Analysis

To validate and further explore the conclusions reported previously,[13] we set out to explore all the computational results using 6-mers of RNA in their single- and double-stranded forms with sequences (**1**-**4**) designed to exist as coils or duplexes in the presence of their complementary strands Figure 1 (top). It is important to note that experimental data of strands containing these sequences displayed the following: 1) formation of structures that hybridized as A-form duplexes; 2) induced thermal stabilization whenever both strands were modified with the methylthiophene units, as recorded through measurement of their thermal denaturation transitions; and 3) overall minor structural changes before and after modification, as recorded via circular dichroism. Consistent with our previous RNA models, strands were modified at the C2'-*O*-position with 2-thiophenylmethyl groups to yield canonical (unmodified) strands **1** and **2** as well as modified strands **3** and **4**, which were modeled as individual RNA coils and their corresponding helical structures.

As displayed on Figure 1 (bottom), side-by-side comparisons between duplexes **1**:**2** and **3**:**4** displayed some visible differences in their overall structures. The main aspect that, at first glance, seemed to be clear was reduced distances between both ends, which is consistent with a compacted structure. To further explore on the nature of these changes, various structural parameters were measured in both single- (**1**-**4**) and double-stranded forms (**1**:**2**, **3**:**4**). Ribose conformational effects can be described using the Altona-Surandalingam parameters, where the phase angle of pseudorotation (*P*) and a puckering amplitude (ψ_m) dictate the form of the duplex, [20] i.e., A-, B-, Z-forms, or polymorphic mixtures containing a range of conformations.[21] Standard conformational isomers of the ribose[22] are represented by these two parameters and can be plotted in a pseudorotational wheel, or individually, by calculating five dihedral angles along the ribose ring (Figure 2A). These

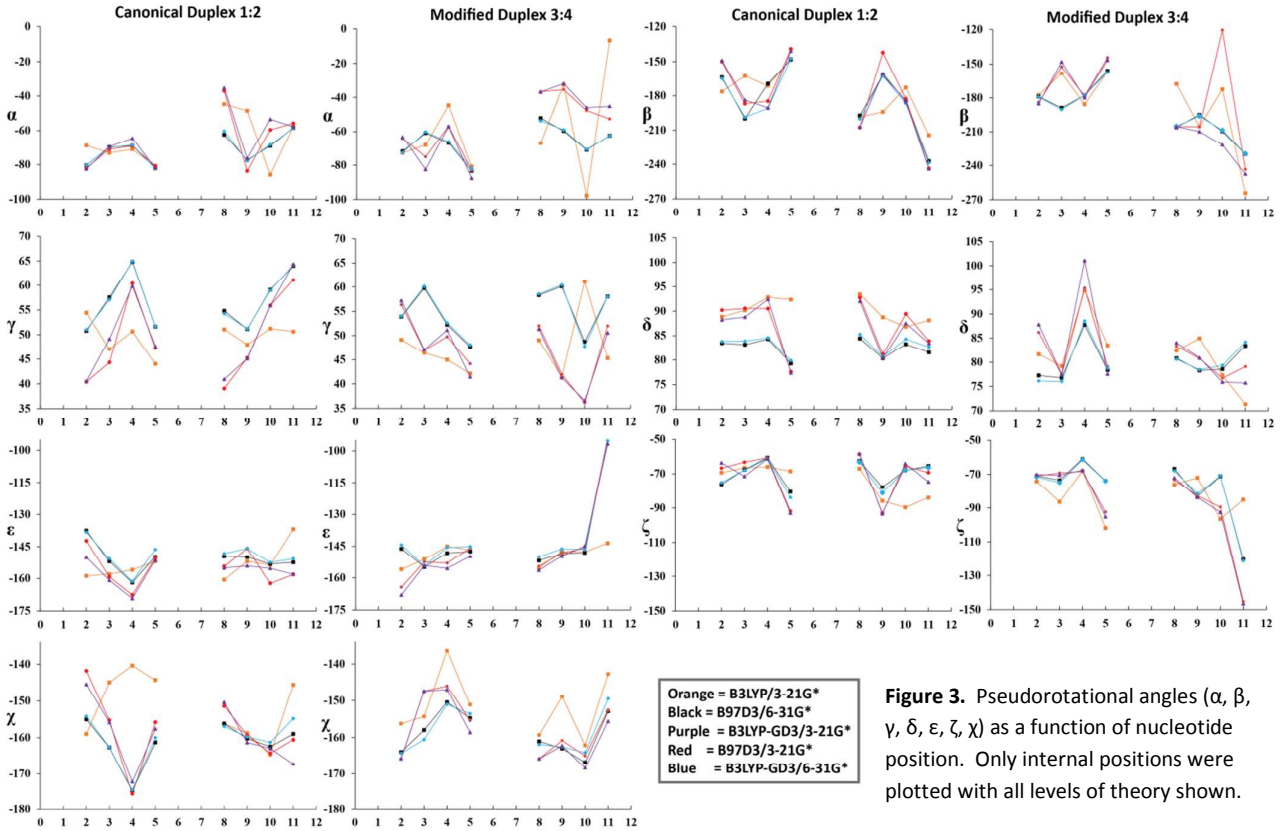

Figure 3. Pseudorotational angles (α, β, γ, δ, ε, ζ, χ) as a function of nucleotide position. Only internal positions were plotted with all levels of theory shown.

relationships can then be plotted as illustrated in Figure 2B, where conformational changes along the ribose is known to occur from the C3'-endo (North) clockwise towards the C2'-endo (South) accordingly. From among the 20 different twist and envelope conformations that are possible,[23] it was found that all coils and duplexes of the canonical and modified strands fell within the range that is expected of an A-form duplex structure. It is worth noting that the single stranded forms of both modified and canonical structures were found

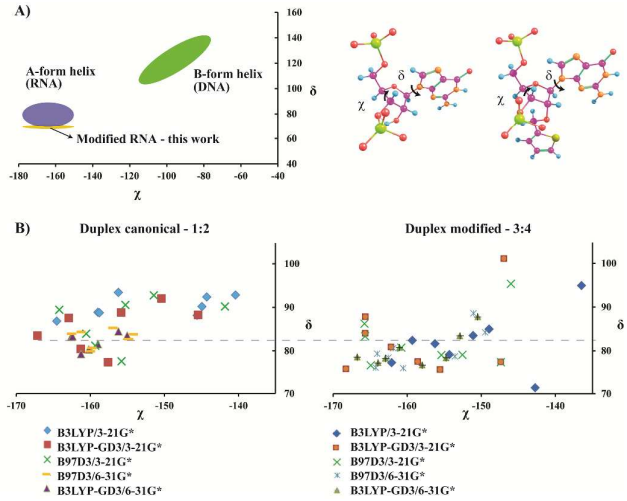

Figure 4. (A) Angle covariance matrix (δ,χ) that highlights the area where B-form and A-form helices are expected (left) along with the modified oligonucleotides presented in this work (top-left). The representation for the dihedral angles plotting the relationship of the nucleobase vs the backbone, with respect to the ribose (top-right). (B) Angle covariance matrix (δ,χ) for canonical **1:2** and modified **3:4** helices. All levels of theory used in this work are plotted.

to have larger variations amongst different positions, which although falling within the expected 0° < P < 36° range, resulted in large standard deviations. The larger variation on single-stranded RNAs is a reasonable observation given the dynamic nature of the biopolymer in this state. Interestingly, calculations carried out on the duplexes resulted in smaller variations and lower overall ranges that corresponded to P values lower than 15°. This observation can be explained due to the higher order imparted by the duplex structure versus the degrees of freedom that are present in coils of the same strands. It is important to note that the dot representation in Figure 2B does not include the range of variation, which was in the order ± 16° compared to a smaller error propagation in the case of the duplexes of ± 4°. The parameters and equations were taken from Altona et al.[24] and are also illustrated in the same figure.

We then proceeded to explore possible changes along the phosphate backbone and the glycosidic bond, which is reflected as a measure arising from torsional angles along these bonds. Nomenclature regarding these dihedral angles (α, β, γ, δ, ε, ζ, and χ) was adopted according to Sundaralingam and references therein (Figure 2C).[25] Gratifyingly, all torsion angles were in close agreement to those reported for a North

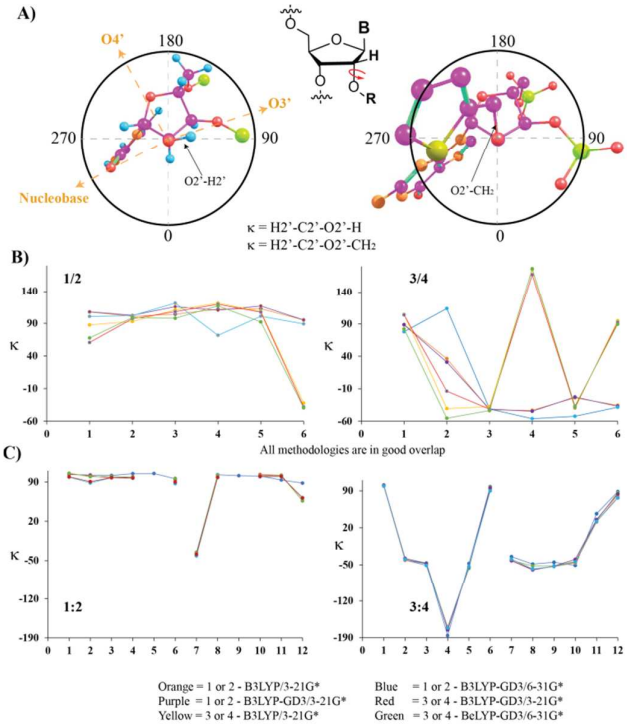

Figure 5. (A) Illustration for the interpretation of the κ angles and their relationship on the C3'-endo (north) conformation. Hydrogen atoms were omitted in the right (modified) structure for clarity. Measured κ angles on (B/C) single-stranded canonical (**1**-**2**) and modified (**3**-**4**) oligonucleotides; and canonical (**1**:**2**) and modified (**3**:**4**) duplexes. Positions 2-5 and 8-11 are functionalized with a thiophenylmethyl group on modified samples.

C3'-endo ribose conformation, also consistent with A-form RNA duplexes.[26] Furthermore, each of the corresponding torsion angles, sugar puckers, and glycosidic torsion angles for the internal residues were analyzed. As shown in the obtained plots (Figure 3), values obtained for canonical and modified oligonucleotides are within the range of an A-form RNA duplex with consistently lower δ values in the case of the modified oligonucleotides. This result is in agreement with a structure that would have a wider and shallower minor groove than that in canonical A-form RNA (11Å) and/or deeper and narrower major groove than typical values of 3 Å, arising from the O4-C4'-C3'-O3 torsional angle. To organize and illustrate the obtained data from these angle,s and track the changes with respect to each other, a representation similar to that previously used to describe DNA:RNA hybrid duplexes was chosen.[27]

As illustrated on the plots in figure 3, no significant discrepancies were observed from calculation of all of the torsional angles (determined as indicated on Figure 2C). All angles fall in the expected region for both canonical and modified cases. With respect to the different computational protocols that were used, a direct comparison between all of them lie within error. One exception for which changes were observed is in the use of B3LYP/3-21G* on both canonical and modified duplexes **1**:**2** and **3**:**4**, with distinct changes on each

of the plots that were not dependent on position, sequence, or identity of the nucleobase. This outlier is most likely due to the deficiency of the functional (B3LYP without dispersion correction) in combination with the smaller basis set (3-21G*). Other results are theoretical more sound and are expected to provide more accurate results. Overall, comparable values were obtained upon comparison of the canonical duplexes with the modified double-stranded samples.

This analysis also shows that the (δ,χ) angle covariance matrices fall within quadrant II, meaning that all δ vs χ plots fall on the expected region of the Cartesian coordinate system corresponding to an A-form right handed antiparallel helical structure.[28] This plot can be seen as a relationship on how the nucleobases are positioned with respect to the structure of the backbone (Figure 4). The orientation of the nucleobase with respect to the ribose ring is described by the glycosidic torsion angle χ_{CN} and was also measured. These data also shows that P-P intrastrand distances of ca. 5.7 Å are within error in both cases and well within the range of the distance for an A-form structure, in contrast with P-P longer distances in a B-form helix, which are expected to be ~ 7 Å apart.[29]

Interestingly, the δ(°) values were consistently lower for the 2-thiophenylmethyl modified duplexes, which suggests a more compacted structure that in turn may explain the reported experimental data displaying higher thermal denaturation transitions, only observed upon modification of both strands. As shown on Figure 4, covariance angles calculated for the modified duplexes were generally lower than a ca. δ 82°, thus providing an explanation for the experimental observations. This result suggests that the small aromatic moieties, while within proximity of each other, are inducing a conformational change along the ribose rings that is acting synergistically from both strands and is stabilizing the structure.

To further explain this observation, we decided to explore the connectivity between the rotational state of the C2'-O2' bond and examine changes in the interactions that appear from having a H-atom versus the methylthiophene functional group. As reported recently,[30] functional groups on this position can act as a conformational switch to tune the corresponding A-form structure. It is known that in the C3'-endo (North) conformation, three direction domains are accessible for the O2'-H2' bond: (i) the O3', (ii) the O4', and (iii) the shallow groove towards the nucleobase (Figure 5).[31] We analyzed the C2'-O2' torsion angle and plotted the obtained values as a function of nucleotide position, which displayed clear differences upon comparison of both the canonical and modified duplexes. As expected, canonical duplexes occupied the space where the O2'-H2' bond is oriented towards the O3' direction with κ angles ranging between 60° - 90°. This is well in agreement with NMR data[32] as well as high-level electronic structure calculations.[33] Interestingly, the κ angle corresponding to the modified oligonucleotides were found in the 310° – 350° range, consistent with the methylene group oriented towards the O4'-direction. Since the direction towards the O3'-position is favored due to H-bonding interactions with water, it makes sense that the presence of a much larger hydrophobic group would change this directionality.

Further analyses of the duplex strands in all levels of theory tested herein resulted in very small discrepancies within each other when comparing amongst modified or canonical strands (Figure 5). Close inspection of the duplex samples indicates that the more pronounced changes occur on nucleotide sites that are modified (positions 2-5 and 8-11), while all measured angles for the canonical duplexes are within range of the expected κ-angles typically measured on A-form RNA. One exception in this respect is position 7, however this can be rationalized as this is the 2'-OH at the C3'-end. Furthermore, κ-angles measured for canonical RNA on single stranded samples did not display significant changes (except for positions 1- and 6- at the ends). This observation differs from modified single stranded samples, where modified positions varied with no distinct pattern. In addition, no clear trend between the levels of theory used was established, presumably due to the inherited randomness of the coils versus that observed in the ordered duplex samples.

To further validate our model while taking into consideration the importance of the 2'-OH, as it has been shown to have an effect on base pair opening and conformational heterogeneity,[34-35] we explored the effect that changes in orientation of this hydroxyl group would have on the overall structure of RNA. We used duplex **1**:**2** to systematically vary the orientation of the C3'-C2'-O2'-O2'H dihedral angle and rotated both strands by 180°, 360°, or one strand 180°/opposing strand randomized angles. It was reasoned that choosing these set of combinations would provide an accurate sense on overall structural changes arising from the different conformational isomers. Each model was optimized and the most stable structures were analyzed in detail. Gratifyingly, it was found that all pseudorotational and torsional angles were in agreement with the previous data thus confirming A-form duplexes (see supporting materials). One angle that displayed slightly higher values was that associated with the κ angle, however, close inspection of the 2'-OH bond does not seem to be involved in unreasonable H-

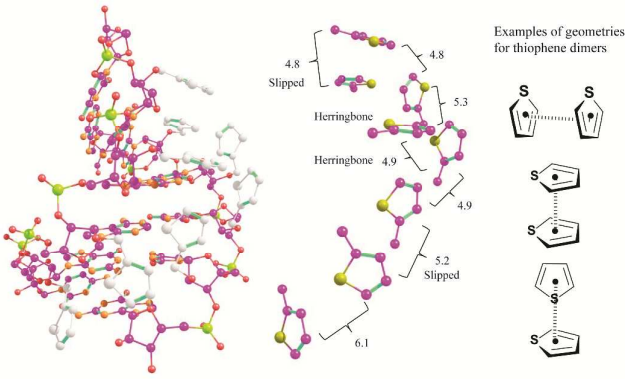

Figure 6. Thiophene rings (colored gray) are shown within the duplex (left) and with the duplex omitted for clarity (center). Centroid-centroid distances (Å) between thiophene rings are shown (right) along with three of at least 20 possible thiophene dimer geometries.

bonding interactions. While changing the orientation of the 2'-OH bond can result in overall geometrical/structural changes[36] this did not seem to affect the overall structure on the duplexes explored herein.

Thiophene π–π-stacking. Another aspect that was considered as a possible contributor to the increased stability on the modified duplexes was π-π interactions between the thiophene groups. The distances from thiophene ring to thiophene ring were obtained by measuring between the centroids on each ring that are in proximity with each other (Figure 6). The obtained centroid-centroid distances ranged between 4.8 and 6.1 Å and are about 1 Å longer than calculated π-stacked rings perfectly aligned as thiophene dimers of ca. 3.8 Å.[37-38] The obtained values however, are more in line with thiophene interactions that include parallel (cofacial), tilted, or herringbone (slipped) arrangements, which measure to be in close agreement with optimized distances of ca. 4.9 Å[39] and that potentially show a variety of different dimer geometries (Figure 6, right). In addition, distances between the sulfur atom and the nearest carbon on adjacent thiophene rings displayed values as low as 3.5 Å; and S•••S distances between 4-6 Å.Overall, while π-stacking interactions may be contributing to the overall stability, it is likely that they are weak. The results suggest that increased stability may be achieved by using molecules with increased energy of aromatization or π-π stacking interactions.

Conclusions

In this work we have performed DFT calculations on 6-mers of RNA in their single- and double-stranded forms. Three functionals, B3LYP, B3LYP-GD3, and B97D3, as well as two basis sets, 3-21G* and 6-31G*, were employed in our study. The solvent environment was modelled using the PCM approach. Structures of all the molecular systems in our study were completely optimized in the electronic structure theory calculations. The consistent prediction of different levels of theory is able to validate the methodology developed on 6-mers of RNA using canonical and modified (2'-*O*-thiophenylmethyl) oligonucleotides. It was noted that to obtain a more complete assessment on the structural parameters of the corresponding coil structures, a higher number of conformational states needs to be sampled and is something that will be addressed in the near future.

In the case of the duplex structures, it was observed that with the canonical structures, all of the measured parameters fell within range of the expected values that are experimentally observed using well established methodology, i.e., NMR, crystallography. Importantly, the results that were obtained from analyses of the modified oligonucleotides aided in the explanation of the experimental results, which indicate that stabilization should occur as long as both of the strands are modified. The measured parameters indicated that in all cases, the (δ, χ) covariance matrix, lies in the expected quadrant II for an A-form duplex but with lower δ values. In agreement with a compacted structure that explains stabilization of these duplexes. It was determined that the nature of this variation is contributed from analysis of the O2'-H2' torsion angle (κ°), in which the 2'-position changes in orientation when the thiophene moiety is present. This in turn induces a synergistic conformational change that results in stabilization of the modified duplex. The effect after varying the orientation of the 2'-OH was also explored and it did not have a major impact on the overall duplex structures. Lastly, π-π stacking interactions were ruled out as a major contributor of increased stability. Minor interactions are possible and are expected to be of higher contribution with larger aromatic systems. Overall the developed methodology should enable future research with accurate prediction of the effect that various modifications will have on duplexes and is guiding our efforts in the design of new functional oligonucleotides of RNA.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work has been supported by the National Science Foundation CHE-1500285 (HW), and used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. M. J. R. E. acknowledges support via start-up funds and a CLAS Research Innovation Seed Program grant from the University of Colorado Denver.

Notes and references

1 Wan, W. B., Seth, P. P. *J. Med. Chem.* 2016, **59**, 9645 – 9667
2 Anderson, B. A.; Onley, J. J.; Hrdlicka, P. J. *J. Org. Chem.* 2015, **80**, 5395-5406.
3 Li, S.; Jiang, Q.; Liu, S.; Zhang, Y.; Tian, Y. et al. *Nat. Biotech*, 2018, **36**, 258-264.
4 Thayer, A. M. *C & EN*. 2017, April **24** issue, pp. 15-17.
5 Mutisya, D.; Hardcastle, T.; Cheruiyot, S. K.; Pallan, P. S.; Kennedy, S. D.; Egli, M.; Kelley, M. L.; Smith, A. van B.; Rozners, E. *Nucleic Acids Res.* 2017, **45**, 8142-8155.
6 Sheng, G.; Gogakos, T.; Want, J.; Zhao, H.; Serganov, A.; Juranek, S.; Tuschl, T.; Patel, D. J.; Wang, Y. *Nucleic Acids Res.* 2017, **45**, 9149-9163.
7 Malek-Adamian, E.; Guenther, D. C.; Matsuda, S.; Martínez-Montero, S.; Zlatev, I.; Harp, J.; Patrascu, M. B.; Foster, D. J.; Fakhoury, J.; Perkins, L.; Moitessier, N.; Manoharan, R. M.; Taneja, N.; Bisbe, A.; Charisse, K.; Maier, M.; Rajeev, K. G.; Egli, M.; Manoharan, M.; Damha, M. J. *J. Am. Chem. Soc.* 2017, **139**, 14542-14555.
8 Košutić, M.; Jud, L.; Da Veiga, C.; Frener, M.; Fauster, K.; Kreutz, C.; Ennifar, E.; Micura, R. *J. Am. Chem. Soc.* 2014, **136**, 6656-6663.
9 Li ,L.; Szostak, J. W. *J. Am. Chem. Soc.* 2014, **136**, 2858-2865.
10 Das, S.; Samanta, P. K.; Pati, S. K. *New J. Chem.* 2015, **39**, 9249-9256.
11 Kohn, W.; Sham, L. J. *Phys. Rev.* 1965, **4A**, A1133-A1138.
12 Deleavey, G. F.; Damha, M. J. *Chem. Biol.* 2012, **19**, 937-954.

13 Nguyen, J. C.; Dzowo, Y. K.; Wolfbrandt, C.; Townsend, J.; Kukatin, S. Wang, H.; Resendiz, M. J. E. *J. Org. Chem.* 2016, **81**, 8947-8958.
14 Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; et al. Gaussian, Inc., Wallingford CT, 2009.
15 Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comp. Chem.* 2011, **32**, 1456-65.
16 Grimme, S.; Antony, J.; Ehrlich, S.;. Krieg, K. *J. Chem. Phys,* 2010*,* **132**, 154104.
17 Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys*. 1997, **107**, 3210-3221.
18 Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A*, 1998, **102**, 870-875.
19 Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A*, 2000, **104**, 4690-4698.
20 Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* 1972, **94**, 8205-8212.
21 Nowotny, M.; Gaidamakov, S. A.; Crouch, R. J.; Yang, W. *Cell*, 2005, **121**, 1005-1016.
22 Lescrinier, E.; Froeyen, M.; Herdewijn, P. *Nucleic Acids Res.* 2003, **31**, 2975-2989.
23 Plavec, J.; Tong, W.; Chattopadhyaya, J. *J. Am. Chem. Soc.* 1993, **115**, 9734-9746.
24 Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* 1973, **95**, 2333-2344.
25 Sundaralingam, M. *J. Am. Chem. Soc.* 1971, **93**, 6644-6647.
26 Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M. in Nucleic Acids in Chemistry and Biology, 3rd ed. Chapter 2 pp. 12-75. RSC 2006. ISBN-10:0-85404-654-2.
27 Fedoroff, O. Y.; Salazar, M.; Reid, B. R. *J. Mol. Biol.* 1993, **233**, 509-523.
28 Anosova, I.; Kowal, E. A.; Dunn, M. R.; Chaput, J. C.; Van Horn, W. D.; Egli, M. *Nucleic Acids Res.* 2016, **44**, 1007-1021.
29 Rich, A. *Nature Struct. Biol.* 2003, **10**, 247-249.
30 Darré, L.; Ivani, I.; Dans, P. D.; Gómez, H.; Hospital, A.; Orozco, M. *J. Am. Chem. Soc.* 2017, **138**, 16355-16363.
31 Auffinger, P.; Westhof, E. *J. Mol. Biol.* 1997, **274**, 54-63.
32 Fohrer, J.; Hennig, M.; Carlomagno, T. *J. Mol. Biol.* 2006, **356**, 280-287.
33 Mládek, A.; Banáš, P.; Jurečka, P.; Otyepka, M.; Zgarbová, M.; Šponer, J. *J. Chem. Theor. Comput.* 2014, **10**, 463-480.
34 Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell Jr., A. D.; *J. Comput. Chem.* 2011, **32**, 1929-1943.
35 Mládek, A.; Banáš, P.; Jurečka, P.; Otyepka, M.; Zgarbová, M.; Šponer, J. *J. Chem. Theory Comput*. 2014, **10**, 463-480.
36 Szabla, R.; Havrila, M.; Kruse, H.; Šponer, J. *J. Phys. Chem. B*, 2016, **120**, 10635-10648.
37 Huber, R. G.; Margreiter, M. A.; Fuchs, J. E.; von Grafenstein, S.; Tautermann, C. S.; Liedl, K. R.; Fox, T. *J. Chm. Inf. Model.* 2014, **54**, 1371-1379.
38 Rodríguez-Ropero, F.; Casanovas, J.; Alemán C. *J. Comput. Chem.* 2008, **29**, 69-78.
39 Tsuzuki, S.; Honda, K.; Azumi, R. *J. Am. Chem. Soc.* 2002, **124**, 12200-12209.