THE JOURNAL OF PHYSICAL CHEMISTRY B

Comment

pubs.acs.org/JPCB

Reply to "Comment on 'Computational Model for Predicting Experimental RNA and DNA Nearest-Neighbor Free Energy Rankings'"

Charles A. Johnson, Richard J. Bloomingdale, Vikram E. Ponnusamy, Conor A. Tillinghast, Brent M. Znosko,* and Michael Lewis*

Department of Chemistry, Saint Louis University, 3501 Laclede Avenue, Saint Louis, Missouri 63103, United States

To start, we agree with Sponer and co-workers[1] that the statement "there most definitely is a quantitative correlation between the quantum mechanical gas-phase stacking data and nucleic acid stability" in our recent paper[2] is too optimistic. Instead, this statement should be tempered to "there most definitely is a quantitative correlation between the *sum of* the quantum mechanical gas-phase stacking *and hydrogen bonding* data and the nucleic acid stability." More important to us, we regret including the numbers of Sponer and co-workers in the Table 3 comparisons in our paper. We misinterpreted the tables in the paper of Sponer and co-workers,[3] and it was not until we read their Comment that we realized the DNA and RNA nearest-neighbor rankings in their work do not include H-bonding and thus should not be compared to our work or to the experimental nearest-neighbor free energy rankings. However, neither the optimistic statement nor the misinterpretation of Sponer's earlier tables alter the major point of our manuscript: the ranking of computational nearest-neighbor energies (composed of hydrogen bonding, interstrand base stacking, and intrastrand base stacking) calculated from average diffraction data do an excellent job of predicting the ranking of experimental nearest-neighbor free energies for RNA and DNA.

Our approach in the original paper was based on our opinion that both hydrogen bonding and base stacking are important to the overall stability of both RNA and DNA. As Sponer and co-workers note in the first point of their Comment, the 10 unique four-nucleotide nearest-neighbor (NN) combinations can be categorized into three categories: the three NN systems containing only AT(U) H-bonding pairs that will have a total of four H-bonds, the four NN systems containing one AT(U) and one GC H-bonding pair that will have a total of five H-bonds, and the three NN systems containing only GC H-bonding pairs that will have a total of six H-bonds. Our work used structures from average fiber diffraction data, and thus all of the AT(U) H-bonding pairs have the exact same structure and H-bonding energy, and the same is true for all GC H-bonding pairs. As a result, the three NN systems with only AT(U) pairs have the exact same H-bonding energies, as do the four NN systems with one AT(U) and one GC pair, and as do the three NN systems with only GC pairs. Figures S1 and S2 from the Supporting Information of Sponer's and co-workers' Comment makes the following point: our H-bonding energies correctly order the three smallest ΔG systems, the four intermediate ΔG systems, and the three largest ΔG systems. Given that six H-bonds are stronger than five H-bonds, which are stronger than four H-bonds, this is exactly what one would expect.

The H-bonding contributions to the overall experimental NN thermodynamic values are dominant, and this is exemplified by the fact that the NN systems with only AT(U) pairs have the three smallest experimental ΔG values, the NN systems with one AT(U) and one GC pair have intermediate ranked ΔG values, and the NN systems with only GC pairs have the three largest ΔG values.[4] Thus, the challenge in ranking the 10 unique NN systems is differentiating among the three AT(U) only systems, among the four AT(U)/GC systems, and among the three GC only systems. Our method of adding interstrand and interstrand base stacking energies to the hydrogen bonding energies was an attempt at differentiating within these three systems. Base stacking energies were able to differentiate within these three systems in some cases. For example, the $E_{NN,Calc\text{-}H}$ values for the three GC only systems for both RNA and DNA (Table 3 in our original paper) were able to rank the three GC systems identical to the ranking of the experimental NN values but were unable to replicate exactly the experimental rankings in all cases.[2] As we noted in our paper, and as Sponer and co-workers reiterate in their Comment, there are obviously many contributing factors that are not accounted for using our method, including entropy effects, hydration, ion binding, sequence-dependent hydrogen bonding energies, etc. Also, the calculations presented in our paper do not consider the fact that the experimental thermodynamics of the duplex were measured with respect to the thermodynamics of the single strands. Nevertheless, our "oversimplified" approach works well enough to predict experimental RNA free energy rankings with a mean absolute difference (MAD) of 1.0 and a Spearman rank correlation coefficient (r_s) of 0.88 and DNA free energy rankings with a MAD value of 0.4 and an r_s value of 0.98 (Table 3 of original paper).[2] If the model was adjusted to include some of the omissions mentioned above, perhaps this could improve the differentiation within systems, the MAD values, and the r_s values. If we, Sponer and co-workers, or others were able to calculate nearest neighbor energies using hydrogen bonding alone, base stacking alone, or a combination of hydrogen bonding and base stacking that resulted in improved MAD values and Spearman coefficients, we believe the results should be treated as important rather than "trivial."

In point 1 of their Comment, Sponer and co-workers also discuss the issue of standard deviation as it pertains to ranking.

Received: April 3, 2012
Revised: May 29, 2012
Published: June 11, 2012

dx.doi.org/10.1021/jp303191v | *J. Phys. Chem. B* 2012, 116, 8333−8334

They point out that some of the experimental nearest neighbor values are within experimental error, specifically noting the RNA experimental nearest neighbor parameters 5′GA3′/3′CU5′ (−2.35 ± 0.06 kcal/mol) and 5′CG3′/3′GC5′ (−2.36 ± 0.09 kcal/mol).[4a] The standard deviations for the 10 unique RNA NN systems range from 0.03 to 0.09 kcal/mol, with an average value of 0.07 kcal/mol.[4a] For the 10 unique DNA NN systems the standard deviation range is 0.04−0.08 kcal/mol and the mean is 0.06.[4b] Taking into account the standard deviation range for each value, 7 of the 10 DNA NN systems have no overlap with the other NN systems. Of the 10 RNA NN systems 5 have standard deviation ranges with no overlap with other NN systems. For DNA, there are two cases where the overlap in standard deviation range is significant enough that it could be reasonably argued the rankings might be incorrect: the 5′CA3′/3′GT5′ and 5′AC3′/3′TG5′ NN systems and the 5′AC3′/3′TG5′ and 5′GA3′/3′CT5′ NN systems. For RNA there are also only two cases where the rankings may be incorrect on the basis of the standard deviation ranges: the 5′GA3′/3′CU5′ and 5′CG3′/3′GC5′ NN systems cited by Sponer and co-workers and the 5′CU3′/3′GA5′ and 5′CA3′/3′GU5′ NN systems. For the other DNA and RNA NN systems that have overlap of the standard deviation ranges, the overlap is minimal and the rankings are likely correct. For simplicity, the standard deviations were ignored in our original paper.[2] Considering the standard deviations and ranking those nearest neighbor parameters within experimental error as equivalent in ranking, the MAD values still remain quite good. For example, when the RNA $E_{NN,Calc\text{-}H}$ ranks are compared to the experimental free energy rankings (when considering standard deviation), the MAD value remains 1.0.

The second point of the Comment by Sponer and co-workers deals with our chosen theoretical methods and the fact that we used fiber diffraction data to obtain our structures. With regard to the latter issue, given the previous strong comments by Sponer and co-workers disparaging the use of average fiber diffraction data for DNA/RNA structures,[3] which we noted in our paper,[2] their comments here are expected. We will obviously have to agree to disagree with Sponer and co-workers. The correlations we report are excellent, and we believe this speaks volumes for the appropriateness of using average diffraction data for starting structures in computational studies involving DNA and RNA nucleic acid bases. With regard to our chosen theoretical method, MP2(full)/6-311G**, we will simply reemphasize the justification we gave in our paper. We acknowledge that this combination of method and basis set does not perform well in reproducing *absolute* arene−arene binding energies obtained at highly correlated methods with complete basis set approximation; however, it does perform very well in reproducing the *relative* arene−arene binding energies at such levels. Our paper reports an approach to predict the experimental nearest-neighbor free energy rankings, and this is inherently a study of *relative* binding energies. As with the point about geometries obtained from average fiber diffraction data, we will simply let the results justify the means.

To finish, we want to point out that Sponer and co-workers Comment is primarily related to the issue of stacking energies and our admitted overreach on the importance of our results in terms of stacking in DNA/RNA bases. However, all issues related to stacking are ancillary to the primary focus of our study, which, as the title suggests, is that computational nearest-neighbor energies, calculated from average diffraction data, do an excellent job of predicting the experimental nearest-neighbor free energy rankings. Thus, although we regret the overreach pertaining to stacking interactions and the misuse of Sponer's previous data[3] in comparisons, the primary findings from our paper remain unchanged.

AUTHOR INFORMATION

Corresponding Author

*E-mail: LewisM5@slu.edu.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Šponer, J.; Morgado, C. A.; Svozil, D. *J. Phys. Chem. B* **2012**, DOI: 10.1021/jp300659f.

(2) Johnson, C. A.; Bloomingdale, R. J.; Ponnusamy, V. E.; Tillinghast, C. A.; Znosko, B. M.; Lewis, M. *J. Phys. Chem. B* **2011**, *115*, 9244−9251.

(3) Svozil, D.; Hobza, P.; Sponer, J. *J. Phys. Chem. B* **2010**, *114*, 1191−1203.

(4) (a) Xia, T.; SantaLucia, J.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. *Biochemistry* **1998**, *34*, 14719−14735. (b) SantaLucia, J., Jr.; Allawi, H. T.; Seneviratne, P. A. *Biochemistry* **1996**, *35*, 3555−3562.