THE JOURNAL OF PHYSICAL CHEMISTRY B

Comment

pubs.acs.org/JPCB

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

Jiří Šponer,*[,†,‡] Claudio A. Morgado,[§] and Daniel Svozil*[,†,∥]

[†]Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic

[‡]CEITEC - Central European Institute of Technology, Masaryk University, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic

[§]Departamento de Química, Universidad Técnica Federico Santa María, Casilla 110-V, Valparaíso, Chile

[∥]Laboratory of Informatics and Chemistry, Faculty of Chemical Technology, Institute of Chemical Technology, Technická 5, 166 28 Prague, Czech Republic

S *Supporting Information*

Recently, Johnson et al.[1] reported interesting quantum-chemical computations on base stacking and base pairing energies in B-DNA and A-RNA. They deduced that "there most definitely is a quantitative correlation between the quantum mechanical gas-phase stacking data and nucleic acid stability". Such conclusion is, in our opinion, too optimistic and oversimplified. We identified several issues questioning the validity of this statement.

1. Johnson et al. reported high correlation between stability orders predicted by computed nearest-neighbor (NN) energies E_{NN} and experimental free energies. However, their E_{NN} energies contain not only stacking but also base pairing (hydrogen bonding) energies. Base pairing and stacking are very different sources of stabilization driven by physically very different principles and, as such, should be treated separately. Yet, from the values of E_{NN} Johnson et al. conclude that there exists a strong correlation between nucleic acid stability and quantum mechanical gas-phase *stacking* data. However, what they actually observe is the effect of hydrogen bonding within individual base pairs. The GC base pair is thermodynamically more stable than the AT/U base pair,[2] which in this particular case correlates with the gas-phase energies.[3] Ten independent stacked base pair steps of canonical NA double helices can be categorized into three groups: three AT/U-only steps, four mixed steps with one GC, and one AT/U base pair, and three GC-only steps with two GC base pairs. The effect of hydrogen bonding is further modulated in a sequence-dependent manner by energy contributions that are yet to be determined. In our preceding studies we concluded[4] that we could not find a good correlation between QM base stacking energy calculations and the experimental free energy order of base pair step stability. In these comparisons, we have intentionally excluded the base pairing energies, because their inclusion creates a "false" and trivial correlation. Indeed, if stacking contributions are omitted from the data, and hydrogen bonding energies are compared to the experimental ΔG°_{37} free energies, very good correlation is found (Figures S1 and S2, Supporting Information). The correlation is evidently driven by base pairing energies that mask the lack of finer correlation for stacking. We and others[4,5] have suggested that in fact the finer free energy effects associated with base sequence may stem from complex balance of numerous energy terms, including not only "intrinsic" base stacking but also entropy effects, hydration, ion binding, etc. Moreover, although stacking calculations are executed with respect to the reference state with no stacking, thermodynamics measurements of duplex formation are done with respect to single strands. Although structural dynamics of single strands is not known, some portion of intrastrand stacking in this reference state can be expected. We wish to point out that the paper by Johnson et al. of course clearly noted that their model could not account for these complex factors.

   In Figures S3−S6 in Supporting Information, we present correlation between the order of base pair step stacking energies in A-RNA and B-DNA derived in our preceding paper,[4b] the net stacking data by Johnson et al.,[1] and the experimental stability order.[2] We think the lack of correlation for B-DNA is quite clear. There is some sign of correlation for A-RNA that reflects the number of AU base pairs, as these have smaller van der Waals area than the GC or AT base pairs and thus stack less efficiently. Still, we think it does not justify to suggest that there is predictive correlation, which would mainly require correlation for base pair steps with identical base pair composition.
2. In addition, ranking by order of ΔG°_{37} values is potentially biasing because it disregards the ∼0.1 kcal/mol experimental errors. Thus, the NN UC/GA (−2.35 kcal/mol) is easily within experimental error of CG/CG (−2.36 kcal/mol). So, they are really the same experimentally. Rather, the correlation should be directly based on the energies, which we show in Figures S7−S10 in Supporting Information. Figures S11 and S12 then clearly show that also including hydrogen bonding energies calculated by us for each base pair step separately from MD simulations[4b] to our stacking

Received: January 19, 2012
Revised: May 31, 2012
Published: June 11, 2012

dx.doi.org/10.1021/jp300659f | *J. Phys. Chem. B* 2012, 116, 8331−8332

energies[4b] introduces correlation with experimental ΔG°_{37} values.

3. Besides that, we also think that the quality of the QM computations by Johnson et al. might be insufficient to guarantee quantitative data. Achieving quantitative accuracy in this kind of computations is anyway not straightforward. For example, in the case of computations by Johnson et al., the MP2/6-311G** computations may underestimate the absolute value of stacking (at least compared to the current reference QM estimated CCSD(T) complete basis set data available in the literature).[4b] The 6-311G** basis set misses any diffuse functions and thus neglects the decisive part of the London dispersion energy. The MP2 method is known to be not sufficiently accurate for base stacking. Indeed, the minimal requirement for stacking calculations is at least one set of diffuse polarization functions,[6] and theory offers many options on how to calculate base stacking better, ranging from the highest accuracy wave function theory approaches up to modern parametrized dispersion-corrected DFT methods.[7] Justification of the MP2/6-311G** method based[8] on benchmarking just for one system ($C_6F_6\cdots C_6H_6$ dimer) is, in our opinion, not fully convincing.

   The possible unfavorable effect of unsuitably chosen basis set is demonstrated by positive values of intrastrand stacking in some of B-DNA and A-RNA base stacks, leading even to positive total base pair step stacking energies in UA/UA (+5.01 kcal/mol) and AA/UU (+2.11 kcal/mol) A-RNA steps (after subtracting the base pairing energies). In our opinion, base stacking should not be in positive energy territory; this may be an indication of some imbalance in the computations, though it is important to tell there do not exist direct experimental data for intrinsic energetics of base stacking.
4. Another source of possible errors biasing reference energy computations is the use of low-resolution fiber diffraction geometries,[4,7a] though we admit that also other choices of geometries for such calculations (including our own approaches) do have limitations.

In conclusion, although we agree that the inclusion of calculated base pairing energies into correlation with experimental thermodynamics data as done by Johnson et al. is per se the correct approach, it then precludes derivation of correlations of stacking energies.

■ ASSOCIATED CONTENT

Supporting Information

Figures showing the relationships between hydrogen bonding energies calculated by Johnson et al.,[1] and the experimental nearest-neighbor (NN) ΔG°_{37} values[2] in A-RNA and B-DNA. Figures showing the relationships between base pair step stacking energies in A-RNA and B-DNA derived in our preceding paper,[4b] the new data by Johnson et al.,[1] and the experimental nearest-neighbor (NN) ΔG°_{37} values.[2] Figures showing the relationships between the sum of stacking and hydrogen bonding energies derived in our preceding paper,[4b] and the experimental nearest-neighbor (NN) ΔG°_{37} values[2] in A-RNA and B-DNA. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: J.S., sponer@ncbr.muni.cz; D.S., svozild@vscht.cz.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by research grant MSM 6046137302, and Grant Agency of the Czech Republic [grant numbers P208/10/2302, 203/09/1476, P208/12/1878, and P305/12/G03]. It was also supported by "CEITEC - Central European Institute of Technology" (CZ.1.05/1.1.00/02.0068) from European Regional Development Fund. C.A.M. gratefully acknowledges the support of FONDECYT (project No11100062) and DGIP-UTFSM (project No131128).

■ REFERENCES

(1) Johnson, C. A.; Bloomingdale, R. J.; Ponnusamy, V. E.; Tillinghast, C. A.; Znosko, B. M.; Lewis, M. *J. Phys. Chem. B* **2011**, *115*, 9244−51.

(2) (a) SantaLucia, J., Jr.; Allawi, H. T.; Seneviratne, P. A. *Biochemistry* **1996**, *35*, 3555−62. (b) Xia, T.; SantaLucia, J., Jr.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. *Biochemistry* **1998**, *37*, 14719−35.

(3) Sponer, J.; Jurecka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142−10151.

(4) (a) Sponer, J.; Jurecka, P.; Marchan, I.; Luque, F. J.; Orozco, M.; Hobza, P. *Chem.—Eur. J.* **2006**, *12*, 2854−65. (b) Svozil, D.; Hobza, P.; Sponer, J. *J. Phys. Chem. B* **2010**, *114*, 1191−203.

(5) (a) Kopitz, H.; Zivkovic, A.; Engels, J. W.; Gohlke, H. *ChemBioChem* **2008**, *9*, 2619−22. (b) Koller, A. N.; Bozilovic, J.; Engels, J. W.; Gohlke, H. *Nucleic Acids Res.* **2010**, *38*, 3133−46. (c) Yildirim, I.; Stern, H. A.; Sponer, J.; Spackova, N.; Turner, D. H. *J. Chem. Theory Comput.* **2009**, *5*, 2088−2100.

(6) (a) Hobza, P.; Sponer, J.; Polasek, M. *J. Am. Chem. Soc.* **1995**, *117*, 792−798. (b) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247−3276.

(7) (a) Sponer, J.; Riley, K. E.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2595−2610. (b) Riley, K. E.; Pitonak, M.; Jurecka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023−5063.

(8) Beg, S.; Waggoner, K.; Ahmad, Y.; Watt, M.; Lewis, M. *Chem. Phys. Lett.* **2008**, *455*, 98−102.