

Water–Nucleobase “Stacking”: H– π and Lone Pair– π Interactions in the Atomic Resolution Crystal Structure of an RNA Pseudoknot

Sanjay Sarkhel,[†] Alexander Rich,[‡] and Martin Egli^{*,†}

Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37235, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received April 24, 2003; E-mail: martin.egli@vanderbilt.edu

Noncovalent interactions assume a significant role in chemistry and biology.^{1,2} While the virtues of the traditional hydrogen bonds, i.e., N–H \cdots O and O–H \cdots O, include key roles in macromolecular structure and function,³ such as overall folding, ligand recognition, hydration, and dynamics, the relatively weaker interactions (as compared to the conventional hydrogen bonds) have received due recognition lately.⁴ Principally, the definition of a hydrogen bond (D–H \cdots A, where D and A are mainly N, O atoms) can be extended to include less electronegative atoms, such as D = C, S and A = π , S. Interactions such as C–H \cdots O,⁵ C–H $\cdots\pi$,⁶ and N–H $\cdots\pi$ ⁷ have been widely studied in macromolecular structures. This then opens up a spectrum of different interactions, from the stronger electrostatic hydrogen bonds to the weaker dispersive interactions.⁸

The hydrogen bond acceptor capability of aromatic rings has been demonstrated both by experimental and theoretical studies.^{7,9} As a prototype, the water \cdots benzene system^{9b,10} has been studied where the permanent quadrupole moment of benzene¹¹ provides substantial negative electrostatic potential for a favorable interaction with the proton(s) of the water molecule. This gives rise to the more commonly observed H– π interaction. By comparison, the interaction of a lone pair directly with the face of the π -system seems counterintuitive and has very little precedence in the literature, mostly in the form of *ab initio* studies on simple model systems with electron-withdrawing groups in the ring.¹² However, crystallographic evidence has been provided for the stabilizing influence of sugar O4' (*l.p.*)– π (nucleobase) intramolecular interactions.¹³

In this communication, we report both H– π and lone pair– π (*l.p.*– π) interactions between water molecules and functionally important unstacked residues in the 1.25 Å crystal structure of the ribosomal frameshifting RNA pseudoknot from beet western yellow virus¹⁴ (Figure 1). Table 1 lists the geometrical parameters for the water– π (*w*– π) interactions. The distances from the ring centroid to the water oxygen atom for A9/W175 and A24/W120, W189 suggest a possible H– π interaction. For C8/W71 and A20/W73, the distances are indicative of *l.p.*– π interactions. The approach of the lone pair is almost perfectly linear for the C8 residue (Table 1, Figure 2a). Recent *ab initio* studies indicate that such an interaction is favorable with electron-withdrawing groups in the ring¹² or a protonated ring system.¹⁵ Scheiner and co-workers have predicted that water prefers the O– π (*l.p.*– π) approach in lieu of the H– π interaction in the case of protonated (ImH⁺) as compared to unprotonated imidazole (Im), respectively.¹⁵ The centroid–oxygen distances reported are 2.9 Å (ImH⁺) and 3.2 Å (Im), and the interaction energies are 8.1 kcal/mol (ImH⁺) and 3.1 kcal/mol (Im). The distances for C8 (2.93 Å) and A20 (3.01 Å) indicate that it is the lone pair of the oxygen atoms that is interacting directly with the face of the π -system (Table 1, Figure 2a,b). The possibility

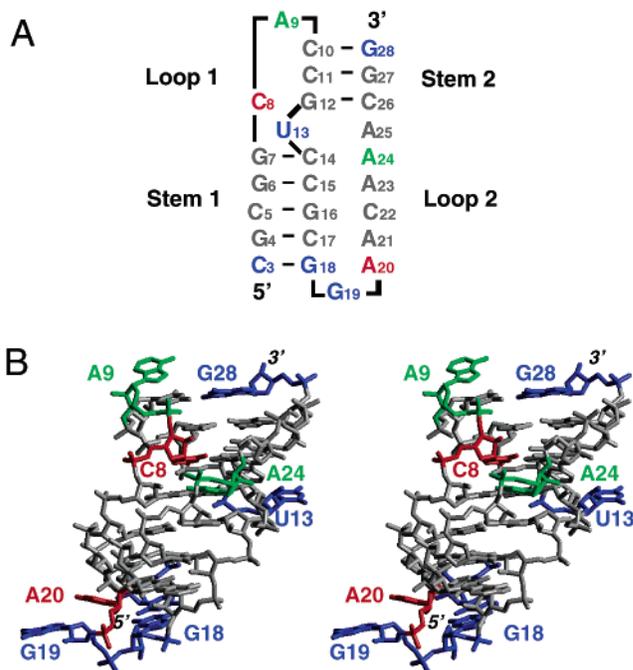


Figure 1. (A) Secondary and (B) tertiary structure (in stereo) of the RNA pseudoknot.¹⁴ Red, green, and blue residues are involved in *l.p.*– π , H– π , and intermolecular base–base stacking interactions, respectively.

Table 1. Geometrical Parameters of Water–Nucleobase Interactions^a

base/water	C*–O _w [Å]	θ [deg]	ω [deg]
C8/W71	2.93	2	91
A20/W73	3.01	10	88
A9/W175	3.40	26	86
A24/W120, W189	3.44, 3.78	7, 18	86, 89

^a C* is the ring centroid, θ is the angle between the normal to the ring plane and the line joining ring centroid and water oxygen (O_w), and ω is the dihedral angle O_w–C*–X₁–X₂, where X₁ and X₂ are the first and second nearest ring atoms to O_w.

of hydrogen atoms interacting with the π -face is unlikely because this would put them at around 2 Å from the centroid, a region that is sterically inaccessible. Furthermore, W71 is hydrogen bonded to the phosphate group of A9 and to W68; most likely the two protons are used as donors in the above cases. The existence of a *l.p.*– π interaction is also consistent with earlier evidence^{16,17} that proved C8 to be protonated at N(3) (see asterisk in Figure 2a). Protonation of C8 possibly reduces the ring electron density and allows a direct interaction with the oxygen lone pair. C8 is engaged in a quadruple base interaction (C8, G12, C26, A25)¹⁶ and is crucial for the frameshifting activity of the pseudoknot as identified by

[†] Vanderbilt University.

[‡] Massachusetts Institute of Technology.

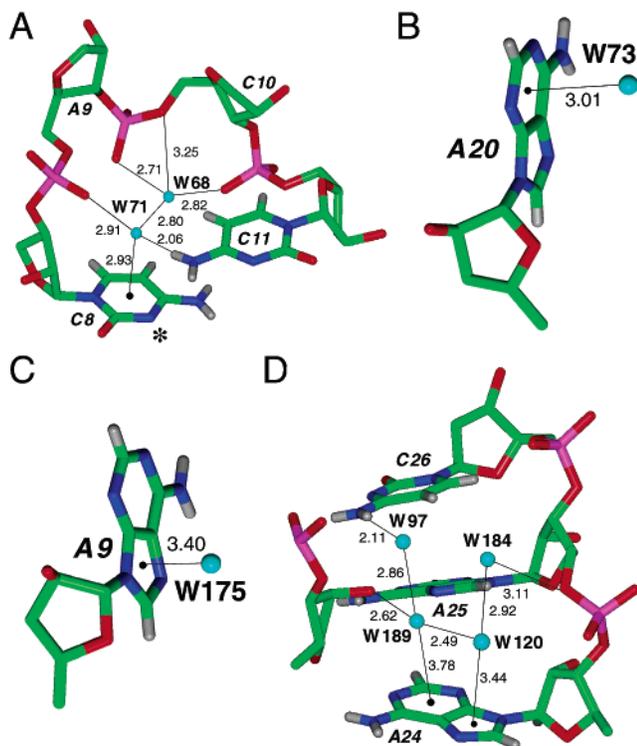


Figure 2.

mutation studies.¹⁸ The additional stabilization provided by the *l.p.*- π interaction may thus be of functional importance.

The *l.p.*- π approach for residue A20 is virtually direct (Table 1, Figure 2b). Interestingly, the water molecule interacts with the six-membered ring rather than the five-membered one. Previous studies on indole found a slightly more favorable H- π interaction with the five-membered ring as compared to the six-membered ring.¹⁵ However, we do note that the electrostatics of the six-membered ring is different in the present case. The electrostatic potential map for adenine shows the centroid region of the six-membered ring to have less electron density.¹⁹ This provides a plausible reason for the observed approach of the lone pair to the face of the π -system. Indeed, the nearest ring atom to the water oxygen atom is C(6), located in a region of positive electrostatic potential.

The interactions of the A9 and A24 residues with water molecules are most likely of the H- π type. W175 interacts with the five-membered ring of A9, the approach being slightly offset (Table 1, Figure 2c). The ring centroid-oxygen distance observed is 3.44 Å. This closely matches the results from earlier *ab initio* studies for such an interaction.¹⁵ The stabilization of the A24 residue presents an interesting situation as W120 and W189 interact with the five- and six-membered rings, respectively. The ring centroid-oxygen distances are 3.44 Å (five-membered) and 3.78 Å (six-membered), and the approaches to the rings are slightly offset (Table 1, Figure 2d). W120 and W189 are hydrogen bonded to each other and are also embedded in a network of H-bonding interactions as shown in Figure 2d. In the above cases, it is interesting to note that the water-nucleobase interactions make up for the absence of base-pair stacking in the looped-out bases. In this regard, the stabilization of the A24 ring by two water molecules is reminiscent of the increased stabilization associated with the enhanced surface

overlap in base-stacking interactions. Classification of such interactions as “water-nucleobase stacking” is therefore justified.

In conclusion, the present study provides a unique example where all RNA residues not involved in either intra- or intermolecular base-base stacking engage in “water-nucleobase stacking”. To the best of our knowledge, this is the first report highlighting *l.p.* (*w*)- π (nucleobase) interactions in an atomic resolution macromolecular structure. Such interactions need to be considered in the analyses of structural data directed at a complete understanding of macromolecular folding and recognition processes. Their inclusion in the parametrization of force fields might also prove to be useful toward accurately representing the role of hydration in the overall stability of macromolecules.

Acknowledgment. Financial support from the National Institutes of Health to M.E. and A.R. is gratefully acknowledged (Grant AI47299).

References

- (1) (a) Lehn, J.-M. *Supramolecular Chemistry: Concepts and Perspectives*; VCH: Weinheim, 1995. (b) Moller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, *100*, 143–167.
- (2) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Oxford University Press: New York, 1999.
- (3) (a) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97–179. (b) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1991.
- (4) Weiss, M. S.; Brandl, M.; Sühnel, J.; Pal, D.; Hilgenfeld, R. *Trends Biochem. Sci.* **2001**, *26*, 521–523.
- (5) (a) Derewenda, Z. S.; Lee, L.; Derewenda, U. *J. Mol. Biol.* **1995**, *252*, 248–262. (b) Wahl, M. C.; Sundaralingam, M. *Trends Biochem. Sci.* **1997**, *22*, 97–102. (c) Ghosh, A.; Bansal, M. *Acta Crystallogr.* **1999**, *D55*, 2005–2012. (d) Jiang, L.; Lai, L. *J. Biol. Chem.* **2002**, *277*, 37732–37740. (e) Pierce, A. C.; Sandretto, K. L.; Bemis, G. W. *Proteins: Struct., Funct. Genet.* **2002**, *49*, 567–576.
- (6) (a) Nishio, M.; Umezawa, Y.; Hirota, M.; Takeuchi, Y. *Tetrahedron* **1995**, *51*, 8665–8671. (b) Nishio, M.; Hirota, M.; Umezawa, Y. *The CH/ π Interaction. Evidence, Nature, and Consequences*; Wiley-VCH: New York, 1998. (c) Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. *J. Mol. Biol.* **2001**, *307*, 357–377.
- (7) (a) Burley, S. K.; Petsko, G. A. *FEBS Lett.* **1986**, *203*, 139–143. (b) Perutz, M. F. *Philos. Trans. R. Soc., Ser. A* **1993**, *345*, 102–112. (c) Worth, G. A.; Wade, R. C. *J. Phys. Chem.* **1995**, *99*, 17473–17482. (d) Worth, G. A.; Nardi, F.; Wade, R. C. *J. Phys. Chem. B* **1998**, *102*, 6260–6272.
- (8) Desiraju, G. R. *Acc. Chem. Res.* **2002**, *35*, 565–573.
- (9) (a) Levitt, M.; Perutz, M. F. *J. Mol. Biol.* **1988**, *201*, 751–754. (b) Suzuki, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. *Science* **1992**, *257*, 942–945. (c) Malone, J. F.; Murray, C. M.; Charlton, M. H.; Docherty, R.; Lavery, A. J. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 3429–3436. (d) Steiner, T.; Koellner, G. *J. Mol. Biol.* **2001**, *305*, 535–557.
- (10) (a) Gregory, J. K.; Clary, D. C. *Mol. Phys.* **1996**, *88*, 33–52. (b) Kim, K. S.; Lee, J. Y.; Choi, H. S.; Kim, J.; Jang, J. H. *Chem. Phys. Lett.* **1997**, *265*, 497–502. (c) Fredericks, S. Y.; Pedulla, J. M.; Jordan, K. D.; Zwier, T. S. *Theor. Chem. Acc.* **1997**, *96*, 51–55. (d) Sorenson, J. M.; Gregory, J. K.; Clary, D. C. *J. Chem. Phys.* **1997**, *106*, 849–863.
- (11) (a) Williams, J. H. *Acc. Chem. Res.* **1993**, *26*, 593–598. (b) Lührer, M.; Bartik, K.; Dejaegere, A.; Bovy, P.; Reisse, J. *Bull. Soc. Chim. Fr.* **1994**, *131*, 603–606.
- (12) (a) Gallivan, J. P.; Dougherty, D. A. *Org. Lett.* **1999**, *1*, 103–105. (b) Tarakeshwar, P.; Kim, K. S. *J. Chem. Phys.* **1999**, *110*, 8501–8512.
- (13) (a) Egli, M.; Gessner, R. V. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 180–184. (b) Berger, I.; Egli, M.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12116–12121. (c) Berger, I.; Egli, M. *Chem. Eur. J.* **1997**, *3*, 1400–1404.
- (14) Egli, M.; Minasov, G.; Su, L.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4302–4307. (The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org [PDB ID code: 1L2X]).
- (15) Scheiner, S.; Kar, T.; Pattanayek, J. *J. Am. Chem. Soc.* **2002**, *124*, 13257–13264.
- (16) Su, L.; Chen, L.; Egli, M.; Berger, J. M.; Rich, A. *Nat. Struct. Biol.* **1999**, *6*, 285–292.
- (17) Nixon, P. L.; Giedroc, D. P. *J. Mol. Biol.* **2000**, *296*, 659–671.
- (18) Kim, Y.-G.; Su, L.; Mass, S.; O'Neill, A.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14234–14239.
- (19) Boehr, D. D.; Farley, A. R.; Wright, G. D.; Cox, J. R. *Chem. Biol.* **2002**, *9*, 1209–1217.

JA0357801