ACS central science

FIRST REACTIONS

Unusual Hydrogen Bonding Patterns and the Role of the Backbone in Nucleic Acid Information Transfer

Steven A. Benner

Foundation for Applied Molecular Evolution, Firebird Biomolecular Sciences LLC, 13709 Progress Boulevard, Alachua, Florida 32615, United States

Remarkably few data were available in 1953 when Watson and Crick proposed that two strands of DNA paired to give an antiparallel double helix, a twisted ladder held together by regularly sized nucleobase pairs (Figure 1).[1] Nevertheless, the model was immediately accepted as "obviously correct".

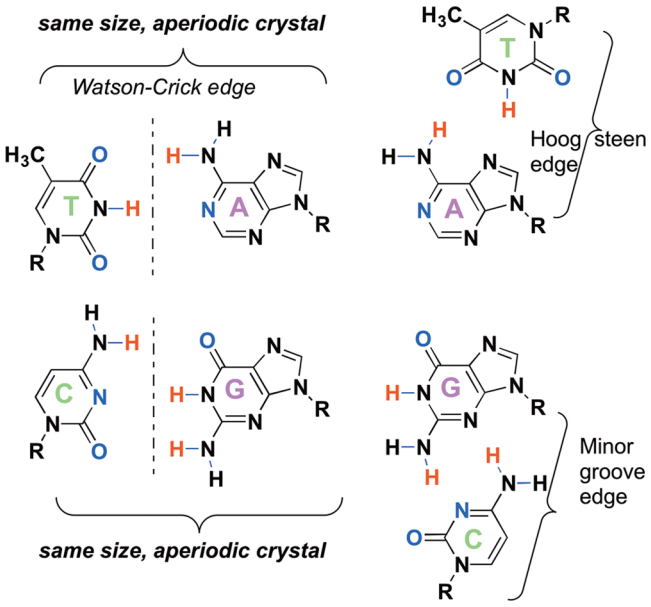

Figure 1. (Left) The "usual" T:A and C:G pairs, whose size similarity allows information to change without disrupting Schrödinger's aperiodic crystal. (Right) The "unusual" pairs, where the pyrimidine approaches the major groove "Hoogsteen" edge of the purine, or the minor groove edge of the purine, fit all of the rules of hydrogen bonding. However, they do not preserve dimensions, and therefore do not support genetics.

Why was this? Culture plays a role in science, as it does in other human activities. Culture is defined by what we accept without explanation, what demands an explanation, and what is "aesthetic". The rapid acceptance of the "elegant" Watson–Crick model can be attributed to the culture of the time.

How so? As it turned out, the regularly sized "rungs" on the two strand duplex ladder fit closely the then-cultural expectation of what a genetic molecule should look like.

Perhaps Watson–Crick nucleobase pairing was not as ubiquitous in the RNA World as we thought?

The rapid acceptance of the "elegant" Watson–Crick model can be attributed to the culture of the time.

These expectations came from a book entitled "What is Life?"[2] Published a decade earlier by Erwin Schrödinger, the book began by noting the thermodynamic "impossibility" of transmitting information at the level at which information was, apparently, transferred by living systems. Schrödinger proposed to resolve this apparent paradox with a genetic polymer that formed an "aperiodic crystal". Here, the overall shape and size of the different information storing units (Schrödinger did not say what they would be) would be the same, so that interchanging the informational bits did not change the overall geometry of the genetic structure.

Watson and Crick immediately recognized that their double helix met Schrödinger's requirement.[3] The size complementarity of the Watson–Crick pair (large purines pair with small pyrimidines) allows information to change without the size of the coding unit changing. This allows the double helix to have the structural features of Schrödinger's aperiodic crystal. The pattern of hydrogen bonding units on the Watson–Crick edge of the nucleobase then ensures that the correct big component finds its small partner.

This was the first reason that the double helix was accepted with little supporting experimental evidence. Second was the "elegance" by which the double helix explained genetic replication. All that was necessary was to separate the two strands, and then allow each strand to

Published: November 22, 2016

DOI: 10.1021/acscentsci.6b00344
ACS Cent. Sci. 2016, 2, 882–884

template the synthesis of its complement using the same pairing rules.

That elegance was taught to generations of students. I can myself remember as a schoolboy being given paper cutouts of A, T, G, and C, as geometric entities to be matched to give equally sized rungs on a ladder, and complementary patterns of hydrogen bond donor and acceptor groups to control specificity.

And it persists today. Many scientists view double helix structure as certainly an optimal, and perhaps a unique, way of solving the information transfer problem in biology, not only on Earth but, perhaps, universally. People are even proposing to replace silicon by DNA as the medium to store massive data, like library collections.

Amid all the hoopla, perplexities embedded in the DNA structure were easy to overlook. For example, adenine is missing one of the three hydrogen bonding units needed as a donor to complement the hydrogen bond acceptor at the 2-position of thymidine. This defect gives DNA a weak nucleobase pair (A:T) and a strong nucleobase pair (G:C), a feature that creates no end of problems to those attempting to do technology with DNA, including creating a set of primers that binds consistently to their targets.

Indeed, when Karst Hoogsteen set out to find evidence for the Watson−Crick pair by cocrystallizing complementary 9-methyladenine and 1-methylthymine, he discovered that they did not interact along the Watson−Crick edge of the heterocycle (Figure 1).[4] Rather, the pyrimidine tucked itself in on the back side of the purine base, forming what is now known as a "Hoogsteen" pair. Here, hydrogen bonding between the two nucleobases satisfies donor and acceptor rules, just not on the right positions.

Many scientists view double helix structure as certainly an optimal, and perhaps a unique, way of solving the information transfer problem in biology, not only on Earth but, perhaps, universally.

Indeed, in the RNA world that likely predated the DNA world, these non-Watson−Crick interactions are well-known.[5] In rRNA, tRNA, and other RNAs that perform nongenetic roles, non-Watson−Crick interactions allow these molecules to fold. And this folding is important, controlling how RNA might have behaved upon emerging in a prebiotic Earth to gain access to the Darwinism necessary to spark the phenomenon that we call "life".

Could these non-Watson−Crick structures also possibly be important for RNA to gain access to Darwinism? If replication involves the addition of single nucleotides, one after another to a growing RNA molecule, some help is certainly needed. The thermodynamics of forming three hydrogen bonds in aqueous solution where water is composed entirely of competing hydrogen bonds seems to be quite inadequate to bind a single nucleotide to a template to form a G:C pair. The situation is worse for the A:T pair, which is joined by just two hydrogen bonds.[6] Various groups have attempted to add downstream units to improve the thermodynamics of monomer association by improving stacking interactions to get monomer addition,[7] but single nucleotide addition is still not a solved problem.

This was the starting point of a paper that appears in *ACS Central Science*.[8] This paper extends efforts in the Szostak laboratory to understand how stepwise template-directed synthesis of RNA might be achieved, absent enzyme. Here, they use using crystallography to observe intermediates in the nucleotide extension process.

To trap those intermediates, Szostak synthesized an analogue of an activated nucleoside phosphate. This was cocrystallized with a rigidified primer−template complex, and a crystal structure was solved. Several types of pairs between G and C were seen. Some were "usual"; they joined G and C in a Watson−Crick geometry. However, other G:C pairs did not join the nucleobases in a Watson−Crick geometry. These were called "unusual". These included structures where the Watson−Crick edge of cytosine was tucked beneath guanine, in the minor groove edge of the purine (Figure 1).

The distinction between the two was cultural, of course, not chemical. In both, hydrogen bonding rules were obeyed; hydrogen bonding complementarity was still seen. And that distinction too was deep in the culture. Citing Leslie Orgel,[9] a hero of the field, Szostak and his co-workers noted that "it has long been thought that the binding of activated monomers to RNA templates would occur predominantly via Watson−Crick base pairing".[8] However, they also note the absence of a reason to exclude "unusual" nucleobase pairs.

But if we step outside of our culture, we might ask the opposite question. Should we not be asking why the "usual" pairs exist? These are the usual pairs that encode the information. What is it about the double helix that enforces nucleobase−nucleobase interactions on the Watson−Crick edge, the interaction that is size uniform, and where the size uniformity fits Schrödinger's requirement for a genetic aperiodic crystal?

DOI: 10.1021/acscentsci.6b00344
ACS Cent. Sci. 2016, 2, 882−884

One feature of the double helix stands out as a possible enforcer of the "usual" Watson−Crick hydrogen bonding interactions: the polyanionic backbone. The repeating backbone phosphate groups make both DNA and RNA acids. Deprotonated at neutral pH, these phosphates dominate the overall physical behavior of nucleic acid molecules. Because of the polyelectrolye backbone, essentially all nucleic acid molecules dissolve in water, precipitate from ethanol, and interact with magnesium, more or less in the same way, regardless of their sequence and the consequent genetic information that those sequences hold.

The remarkable constancy of the physical behavior tied to this molecular structure is also important for Darwinism, not the least because it enforces the aperiodic crystal structure required by Schrödinger. Here, Coulombic repulsion between the backbone charges forces strand−strand interactions as far away from the backbone as possible: the Watson−Crick edge of the nucleobases satisfies this requirement. These observations have led to the "polyelectrolyte theory of the gene", proposed to be a universal guide for the search for life in the cosmos.[10]

> But if we step outside of our culture, we might ask the opposite question. Should we not be asking why the "usual" pairs exist?

Nonionic analogues of DNA and RNA are not held to such a geometry. For example, replacing the anionic phosphate diester linker with the similarly shaped, but uncharged, dimethylenesulfone linker generates DNA and RNA analogues[11] that also give non-Watson−Crick structures.[12] The non-Watson−Crick structures by Zhang et al. are associated with low charge density as well.

Even if the "polyelectrolyte theory of the gene" explains the "usual" Watson−Crick parirs, it is worrisome for those interested in the origins of life. If all hydrogen bonding interactions are more or less equivalent energetically, and polymerases were not present to ensure that only information-important interactions occurred, we encounter a real problem to get the first genetic material copied. The Szostak laboratory has found a noncanonical interaction that is productive. The polyelectrolyte guarantee on information-productive nucleobase−nucleobase interactions only works with a certain number of negative charges preassembled. Getting these is not an insurmountable task for the prebiotic chemist, but one worthy of further examination.

Author Information
E-mail: sbenner@ffame.org.

ORCID
Steven A. Benner: 0000-0002-3318-9917

ACKNOWLEDGMENTS

I am indebted to the Templeton World Charity Foundation and NASA (MCB-1412869) for supporting work behind these ideas.

REFERENCES

(1) Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737−738.
(2) Schrödinger, E. *Was ist Leben*; Serie Piper *1134*; Piper: 1943.
(3) Olby, R. *The Path to the Double Helix: The discovery of DNA*; University of Washington Press: Seattle, 1974.
(4) Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1963**, *16*, 907−916.
(5) Lilley, D. M. J.; Eckstein, F. *Ribozymes and RNA Catalysis*; RSC Publishing: 2008.
(6) Kervio, E.; Claasen, B.; Steiner, U. E.; Richert, C. The strength of the template effect attracting nucleotides to naked DNA. *Nucleic Acids Res.* **2014**, *42*, 7409−7420.
(7) Vogel, S. R.; Deck, C.; Richert, C. Accelerating chemical replication steps of RNA involving activated ribonucleotides and downstream-binding elements. *Chem. Commun.* **2005**, 4922−4924.
(8) Zhang, W.; Tam, C. P.; Wang, J.; Szostak, J. W. Unusual Base-pairing Interactions in Monomer-Template Complexes. *ACS Cent. Sci.* **2016**, DOI: 10.1021/acscentsci.6b00278.
(9) Orgel, L. E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99−123.
(10) Benner, S. A.; Hutter, D. Phosphates, DNA, and the search for nonterrean life: A second generation model for genetic molecules. *Bioorg. Chem.* **2002**, *30*, 62−80.
(11) Richert, C.; Roughton, A. L.; Benner, S. A. Nonionic analogs of RNA with dimethylene sulfone bridges. *J. Am. Chem. Soc.* **1996**, *118*, 4518−4531.
(12) Steinbeck, C.; Richert, C. The role of ionic backbones in RNA structure: An unusually stable non-Watson-Crick duplex of a nonionic analog in an apolar medium. *J. Am. Chem. Soc.* **1998**, *120*, 11576−11580.

DOI: 10.1021/acscentsci.6b00344
ACS Cent. Sci. 2016, 2, 882−884