

Opinion Article

THE DOUBLE HELIX IN RETROSPECT: A CRITIQUE OF AVAILABLE EVIDENCE

Abstract – This short review commentary deals with the evidence for the DNA double helix and raises the intriguing question: does it exist to any significant extent in the solution phase? Well past the fiftieth year of its discovery via X-ray diffraction, the DNA double helix has now acquired the status of a sacred truth, with a large area of modern science revolving around it. This paper presents the pros and cons of the double helix from a physico-chemical perspective, although with critical inputs from chemical biology as appropriate.

Keywords – DNA melting; duplex; Gibbs free energy; helicase; RNA; single strand

INTRODUCTION

The DNA double helix is central to modern biology, perhaps even a sacrosanct principle of science [1-4]! Indeed, the double helix serves to bridge the gulf between classical Mendelian genetics and the modern era of ‘designer genes’. Yet, despite this conceptual inheritance (sic), molecular DNA must conform to the laws of physico-chemical science.

The double helical structure of DNA, possessing as it does an eerie symmetry, is beautiful to behold. It rests on early crystallographic work, the culmination of a brilliant piece of scientific investigation which is as intriguing as a le Carré thriller [4]! Interestingly, however, despite its current status as a bed-rock, the double helix is apparently not beyond the reach of a devil’s advocate, given the complexity of the molecule and the reactions it undergoes. Indeed, a particularly intriguing question relates to the extent of the double helical form in solution – assumed as self-evidently total so far.

Such doubts would be trivial – even cynical – in the case of smaller (and lesser) molecules. However, DNA is far too critically important to be treated dismissively and peremptorily, and indeed far too complex to be buried in blasé assumptions! This paper thus undertakes the onerous task of a careful and dispassionate consideration of the evidence, garnered by some of the most astute scientific investigators over half a century and more.

DISCUSSION

An a priori analysis of the pros and cons of the double helix

Thermodynamic considerations

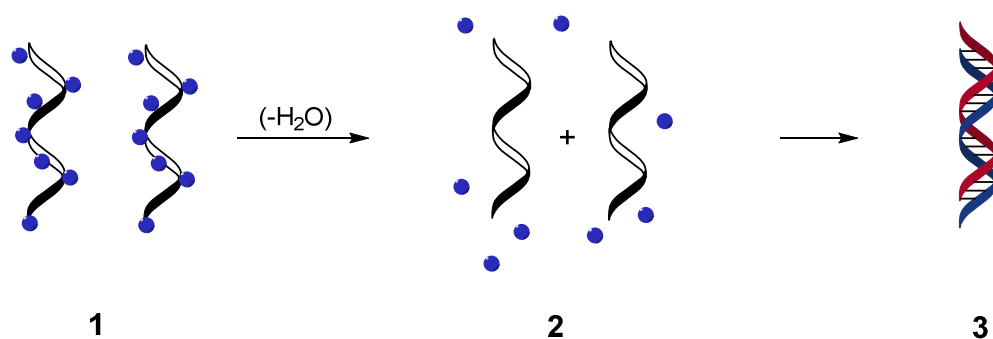
The physico-chemical basis of the double helical form of DNA may be analyzed fundamentally, a thermodynamic approach leading to revealing insights (Scheme 1 and Eqs. 1-6) [5,6]. Thus, the formation of the double-stranded form (*ds*-DNA, **3**) from the corresponding complementary single strands (*ss*-DNA, **1**) would – by implied assumption – be accompanied by a loss of both enthalpy and entropy (*cf.* ΔG_{3-1}). On the other hand, *ss*-DNA would be highly solvated in the aqueous medium (*cf.* **1**), so desolvation (to form **2**) prior to duplex formation (**3**) would be accompanied by gains in both enthalpy and entropy (*cf.* ΔG_{2-1}). The relative extents of these losses and gains would determine the overall stability of *ds*-DNA relative to *ss*-DNA (*cf.* ΔG_{3-1}).

$$G_n = H_n + TS_n \quad (1)$$

$$\Delta G_n = \Delta H_n + T\Delta S_n \quad (2)$$

$$\Delta G_{2-1} = G_2 - G_1 \quad (3)$$

$$\Delta G_{3-2} = G_3 - G_2 \quad (4)$$



Scheme 1. Cartoon representation of DNA duplex formation: **1** is *ss*-DNA that is hydrated at the bases; **2** is *ss*-DNA with its bases stripped of water; **3** is *ds*-DNA. The blue spheres represent water molecules, bound to the bases in **1**, and free in **2**

$$\Delta G_{3-1} = G_3 - G_1 \quad (5)$$

Eqs. 1-5 refer to the Gibbs free energy (G), enthalpy (H) and entropy (S) contents of *ss*-DNA (**1** and **2**) and *ds*-DNA (**3**). (These are hydrated at the bases in **1** but not in **2** and **3**; hydration elsewhere in the molecule is assumed to be similar in **1-3**, hence ignored.) The subscripts ($n = 1-3$) indicate the three states (**1**, **2** and **3**), the corresponding changes being ΔG , ΔH and ΔS . Notably, the duplex is stabilized by hydrogen bonding and base stacking interactions, and destabilized by electrostatic repulsion between the negatively charged sugar-phosphate backbones (but weakly, as these are mutually remote); all these are essentially enthalpic effects, but the duplex would also be destabilized entropically relative to the corresponding single strands.

Thus, the stability of *ds*-DNA over *ss*-DNA depends critically on (ΔG_{3-1}) being substantially negative. This implies (Eq. 6) that the apparent gain in Gibbs free energy upon the desolvation of *ss*-DNA (ΔG_{2-1}) is overwhelmed by the loss of Gibbs free energy upon duplex formation (ΔG_{3-2}): clearly, assumed by current dogma as inviolable.

$$\Delta G_{3-1} = (\Delta G_{2-1} + \Delta G_{3-2}) \ll 0 \quad (6)$$

64 However, it is not easy to assess – even qualitatively – the above losses and gains. Thus, is
 65 the gain in (essentially) solvent entropy upon desolvation of the *ss* form (ΔS_{2-1}) greater or less
 66 than the loss of entropy upon duplex formation (ΔS_{3-2})? Is the enthalpy gain upon the
 67 desolvation (ΔH_{2-1}) greater or less than the enthalpy loss upon duplex formation (ΔH_{3-2})?
 68 (There is also the possibility of enthalpy-entropy compensation in these steps.)

69 *Crystallographic caveats*

70 The existence of crystalline *ds*-DNA is, of course, what is to be expected from considerations
 71 of symmetry and stability. (In other words, the stable *crystalline* form of DNA would be *ds*-
 72 DNA.) Thus, it is possible that *ds*-DNA crystallizes out of a mixture of forms in solution,
 73 even if *ds*-DNA were to be a minuscule component of the mixture, dragging the equilibrium
 74 (between *ss* and *ds* forms) along in the process. This enduring caveat of X-ray
 75 crystallography, understandably acquires particular piquancy in the case of DNA structure!

76 **Experimental evidence for the double helix in solution**

77 *General considerations*

78 Current molecular biology, on one hand, and belief in the existence of *ds*-DNA in solution,
 79 on other, apparently share a symbiotic relationship, reinforced by the ideas of base specificity
 80 and complementarity. In particular, the replication of DNA occurs upon a template strand of
 81 *ss*-DNA, thus bolstering the belief that this leads to *ds*-DNA. Furthermore, the idea of a
 82 compact symmetrical double helical form is not only aesthetically appealing, but also
 83 represents (however inaccurately) a “Mendelian vestige” – that of shared inheritance, *i.e.* one
 84 gene from each parent for every phenotypical characteristic.

85 None of these considerations *per se*, however, proves or disproves the existence of *ds*-DNA
 86 in solution. Thus, DNA replication could occur upon a template strand without prejudice to
 87 the possibility that the *ss* form exists – even exclusively – in solution. The template strand

88 would then be akin to a cofactor for DNA polymerase [7]. (In other words, base pairing
89 occurs only at the active site of the polymerase, with the strands separating on either side of
90 the active site.) One balks at the idea that complementary *ss* forms could thus ignore each
91 other in solution, but why not if thermodynamic considerations deem it so?

92 Thus, the existence of complementary strands could reflect the need for a template at the
93 active site of DNA polymerase, eventually resulting in genetic information being stored in
94 two complementary sets. In other words, if *ss*-DNA can survive on its own, so be it. Although
95 it is intuitively appealing to imagine that in *ds*-DNA the complementary strands somehow
96 protect each other (say, against mutations), this *per se* does not prove that *ds*-DNA
97 predominates in solution.

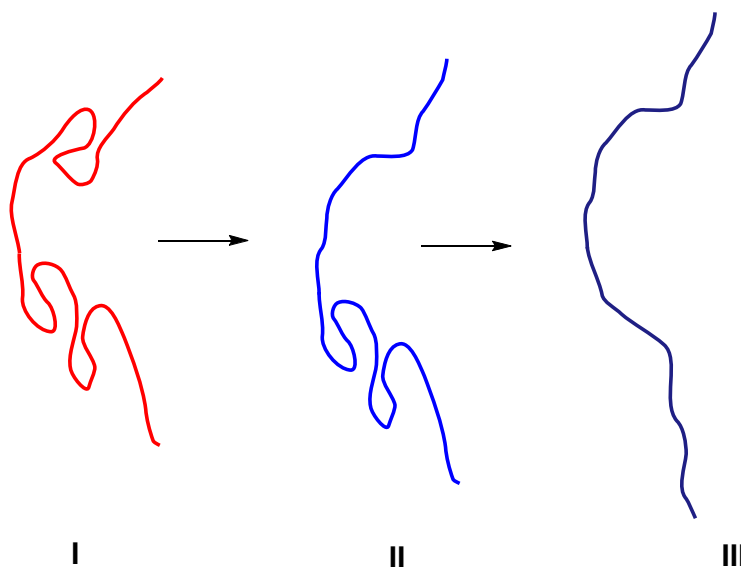
98 *DNA melting and related studies*

99 DNA melting. DNA melting is an intriguing phenomenon which apparently implicates *ds*-
100 DNA [8,9]. Essentially, the ‘melting’ involves heating a solution of DNA while following the
101 change in its ultraviolet absorbance. It is observed that the intensity of the absorption band at
102 ~ 260 nm increases with the temperature, the resulting plot resembling a sigmoid. This is
103 interpreted in terms of the gradual collapse of the duplex structure of *ds*-DNA, with the high
104 temperature plateau region representing complete strand separation to form *ss*-DNA.
105 Accordingly, the mid-point of the sigmoid is considered to represent a 1:1 mixture of *ds*-
106 DNA and *ss*-DNA, the ‘melting temperature’ (T_m) at the mid-point being characteristic of the
107 sample of DNA as defined by its base composition and sequence. (The observed changes can
108 also be reversed, albeit slowly, upon cooling.)

109 This analysis, however, clearly assumes that the initial sample is *ds*-DNA. In fact, all that is
110 being observed is a reversible change, which most likely involves breaking and reforming
111 hydrogen bonds. In fact, a particular problem is that the above analysis assumes a continuous

112 and rapidly attained equilibrium between *ds* and *ss* forms at each temperature: however, it is
 113 known that the (putative) reverse ‘annealing’ process is relatively slow, so reversibility *en*
 114 *route* cannot be assumed.

115 An alternative interpretation to the inter-conversion between *ds*-DNA and *ss*-DNA, however
 116 heretically, would involve only *ss*-DNA. The structure of *ss*-DNA would expectedly be self-
 117 coiled *via* internal base-pairing, akin to that of ribonucleic acid (RNA) [1-3]. Thus, the
 118 ‘melting’ phenomenon would represent the uncoiling of *ss*-DNA leading to various linear
 119 forms consisting of greatly reduced hydrogen bonded base pairs (Scheme 2). (In fact, such
 120 behaviour is indeed known in the case of RNA – famously single stranded [10,11]!)



121 **Scheme 2.** Cartoon representation of the uncoiling of *ss*-DNA, possibly during DNA melting.
 122 **I** represents the fully coiled form, **II** the partly coiled form and **III** the fully extended form.
 123 The transition between **I** and **II**, as also between **II** and **III**, is relatively fast and reversible;
 124 however, the transition from **III** to **I** would be relatively slow. Hence, the overall
 125 transformation would be slowly reversed (‘annealing’), although intermediate states may well
 126 attain rapid equilibrium among themselves.

“Annealing” would then refer to the recoiling of the extended linear forms. This also explains the van’t Hoff plots obtained from DNA melting experiments [8]: the two-state model would still apply, but to an equilibrium between the coiled and extended *ss* forms, rather than between *ds* and *ss* forms. It is also easier to accept the notion of reversibility of the uncoiling of *ss*-DNA, as this would be a fast unimolecular process. The relative slowness of the ‘annealing’ process would then arise from the fact that many *ss* coiled forms are possible, with the intermediate states, however, involving structurally-localized rapid equilibria (Scheme 2).

In fact, an intriguing problem with the current view of DNA melting is noteworthy: the melting of particularly long duplexes would require prohibitively large inputs of energy. Thus, assuming (conservatively) an average base-pairing strength of ~ 5 kcal/mole, even a moderately long 100-mer duplex would involve ~ 500 kcal/mole for complete strand separation. Then, a free energy of activation of at least this magnitude is also indicated. This implies an impossibly slow process, certainly ruled out entirely in the case of naturally occurring DNA!

Related studies. There have been several other elegant and rigorous studies that have been reported on the kinetics and thermodynamics of duplex formation [12-14]. However, these appear for the most part to deal with relatively short oligonucleotides, and the validity of extending the results to the case of naturally occurring DNA is unclear. Shorter oligonucleotide chains cannot form the U-shaped folded structures that are accessible to longer (single stranded) chains, so the shorter chains are predisposed to (intermolecular) duplex formation.

An ingenious strategy involving effective molarity (EM) values for different inter-strand cross-linking reactions [14], generally revealed low EM values (0.8-25) even in the case of moderately long oligonucleotides (19-35 units). (EM’s are a quantitative measure of the

proximity effect, and apparently indicate the possibility of duplex formation.) However, it should be noted that an $EM > 1$ can also arise in the case to two complementary single strands, in which the transition state for the cross-linking reaction is stabilized by base pairing. In fact, when the cross-linking reaction was reversible, the thermodynamic EM's were lower than the kinetic EM's: this possibly indicates that duplex formation is temporary and localized around the reaction centre (*i.e.* the transition state).

Helicases

These are a class of ATP-dependent enzymes that are believed to catalyze the unwinding of *ds*-DNA, leading to both the uncoiling of the double helix and partial strand separation [15]. This process, clearly, must precede certain key reactions of *ds*-DNA, *e.g.* replication, transcription, etc. However, a possibility that needs to be considered is that helicases act upon minor amount of *ds*-DNA present in solution, for a maximal and efficient utilization of the total DNA present. (This would also be an evolutionary justification for the existence of the helicases.) Thus, the mere existence of the helicases does not prove that all of the solution state DNA is *ds*-DNA. These caveats gain credence from the fact that helicases specific to RNA are indeed also known [15], and catalyze the uncoiling of local duplex structures in RNA. Clearly, a similar function in *ss*-DNA cannot entirely be ruled out!

The case of RNA

An enduring question, perhaps obliquely relevant to the present discussion, concerns the structure of RNA in its varied forms. This is almost always single stranded [1-3], so the question is: why not double stranded? Conversely, if RNA is single stranded, why not DNA too? Structurally, it is possible that the C_2 hydroxyl in ribose would sterically hinder duplex formation in the case of RNA, although the extent of this putative interaction appears unknown.

In fact, evolutionary theory posits the existence of a prebiotic RNA world [16]. Interestingly, this raises the question whether duplex formation in DNA is necessary to stabilize it, if single stranded RNA could survive the particularly harsh prebiotic conditions of the primordial world!

CONCLUSIONS

Available evidence is apparently inconclusive about the stable form of DNA – whether single stranded or duplex – in solution. Practically all studies conducted so far have been based on the assumption that the early crystallographic duplex structure may be extended to the solution phase. Thus, the results of DNA melting and related studies are apparently ambiguous, and may equally be interpreted in terms of the uncoiling of single-stranded DNA. In like vein, the helicases may also act upon localised duplex structures in single-stranded DNA (as is known in the case of RNA).

In fact, fundamental thermodynamic considerations indicate that the duplex structure may be unviable. This is because of the prohibitively high energetic cost of separating a putative duplex form of DNA, prior to the plethora of reactions critical to survival (replication, translation, etc.). Thus, intriguingly, the co-existence of complementary strands in solution could be “artefactual”, deriving from the need for a template strand in the action of DNA polymerase (hence does not prove the duplex structure *per se*).

Clearly, further studies are indicated, current evidence being too finely balanced to resolve the dilemma – although perhaps leaning toward single-stranded DNA (in this author’s modest opinion)!

REFERENCES

1. Nelson DL, Cox MM. Lehninger principles of biochemistry. 4th ed. New York: WH Freeman; 2007.

- 201 2. Watson JD, Baker TA, Bell SP, Gann A, Levine M, Losick R. Molecular biology of the
202 gene. 7th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2014.
- 203 3. Allison LA. Fundamental molecular biology. 2nd ed. Hoboken (NJ): John Wiley; 2012.
- 204 4. Watson JD. The double helix: a personal account of the discovery of the structure of DNA.
205 New York: Atheneum; 1968.
- 206 5. Tinoco I Jr, Sauer K, Wang JC, Puglisi JD. Physical chemistry, principles and applications
207 in biological science. 4th ed. Upper Saddle River (NJ): Prentice Hall; 2003.
- 208 6. Nelson P. Biological physics; energy, information, life. New York: WH Freeman; 2004.
- 209 7. Kool ET. Active site tightness and substrate fit in DNA replication. *Annu Rev Biochem.*
210 2002;71:191-219. (doi: 10.1146/annurev.biochem.71.110601.135453).
- 211 8. Howard KP. Thermodynamics of DNA duplex formation. *J Chem Educ.*
212 2000;77(11):1469-71.
- 213 9. Bloomfield VA, Crothers DM, Tinoco I Jr. Nucleic acids: structures, properties and
214 functions. Sausalito (CA): University Science Books; 2000. (pp 101-180).
- 215 10. Puglisi JD, Tinoco I Jr. Absorbance melting curves of RNA. *Methods Enzymol.*
216 1989;180:304-25.
- 217 11. Nikolova EN, Al-Hashimi HM. Thermodynamics of RNA melting, one base pair at a
218 time. *RNA.* 2010;16:1687-91.
- 219 12. Rauzan B, McMichael E, Cave R, Sevcik LR, Ostrosky K, Whitman E, et al. Kinetics and
220 thermodynamics of DNA, RNA and hybrid duplex formation. *Biochemistry.* 2013;52:765-72.
221 (dx.doi.org/10.1021/bi3013005).

- 222 13. Dupuis NF, Holmstrom ED, Nesbitt DJ. Single molecule kinetics reveal cation-promoted
223 duplex formation through ordering of single-stranded helices. *Biophys J.* 2013;105:756-66.
224 ([dx.doi.org/10.1016/j.bpj.2013.05.061](https://doi.org/10.1016/j.bpj.2013.05.061)).
- 225 14. Catalano, MJ, Price NE, Gates KS. Effective molarity in a nucleic acid-controlled
226 reaction. *Bioorg Med Chem Lett.* 2016;26:2627-30.
227 ([dx.doi.org/10.1016/j.bmcl.2016.04.022](https://doi.org/10.1016/j.bmcl.2016.04.022)).
- 228 15. Singleton MR, Dillingham MS, Wigley DB. Structure and mechanism of helicases and
229 nucleic acid translocases. *Annu Rev Biochem.* 2007;76:23–50. (doi:
230 10.1146/annurev.biochem.76.052305.115300).
- 231 16. Neveu M, Kim H-J, Benner SA. The “strong” RNA world hypothesis: fifty years old.
232 *Astrobiology.* 2013;13(4):391-403. (doi:10.1089/ast.2012.0868).