

H-bonds and DNA

Stephen Harding

(University of Nottingham, UK)

Hydrogen bonds or 'H-bonds' are polar, non-covalent bonds or interactions between a hydrogen atom (H) attached to a more electronegative atom, such as oxygen (O) or nitrogen (N), which partially pulls the electron cloud away from the H, leaving it electropositive—with another electronegative atom, such as O or N from a different molecule or from a different part of the same molecule. H-bond interactions play a huge role in the biochemistry of living processes, and in the structures and interactions of biological molecules, with each other and with different molecules including water. Nature's natural solvent, water, is itself a dynamic H-bonded polar structure, which strongly affects solubility and, as (dynamic) water of hydration, interactions between molecules. Compared with covalent and ionic bonds, H-bonds are individually much weaker (<20 kJ/mol), which make them ideal for molecular recognition phenomena. When many H-bonds come together they can form strong insoluble structures such as cellulose and the impermeable derivative of cellulose known as chitin, or helical structures with intra-chain stabilizing H-bonds such as the α -helix. Perhaps the most important H-bonded structure of them all is DNA.

In 1947, the discovery of H-bonds in DNA proved to be one of the crucial steps in the final elucidation of its structure six years later, brilliantly solved by Watson and Crick who were able to put all the experimental pieces of the scientific jigsaw together to yield the double-helical structure. The H-bonds between the pyrimidine and purine bases of each chain (Figure 1) were shown to be strong enough to hold the two chains together in the intact double helical structure but weak enough to allow the chains to be zipped apart for replication.

Discovery of H-bonds in DNA

The discovery of H-bonds in DNA is an interesting story. It was the culmination of the foresight and hard work of a number of researchers and involved meticulous attention to sample purity and integrity, which proved critical for the physico-chemical characterizations. The work

commenced in the 1940s by a team of researchers at (what was then) University College Nottingham—an external college of the University of London. This was before the University of Nottingham received its Royal Charter as a University in 1948. The team was led by Professor John Masson Gulland, the Jesse Boot Professor of Chemistry and assisted by Dr Denis O. 'Doj' Jordan and a succession of three PhD students – Cedric Threlfall, H.F.W. Taylor and J. Michael Creeth. The results were released in a series of three papers all published in 1947 in the *Journal of the Chemical Society*. The first paper described the production of high-purity, non-degraded high molecular mass DNA, the second reported the acid and alkali titration studies on this high-purity DNA and the final and definitive paper identifying H-bonding reported the effects of increasing amounts of acid or alkali on the viscosity and streaming birefringence properties of the DNA at high and low pH values.

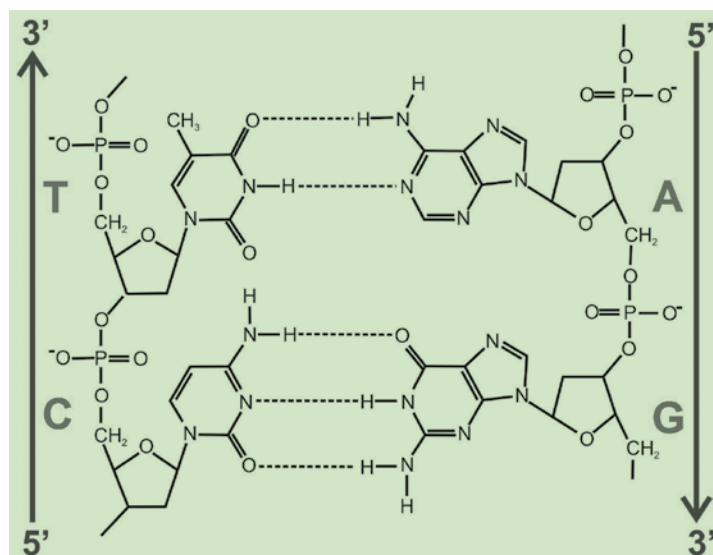
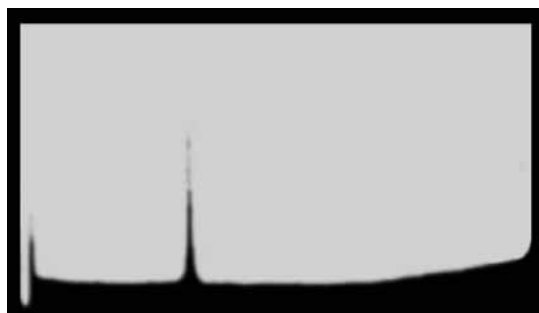


Figure 1. H-bonds in DNA. Part of the Watson and Crick DNA structure with pyrimidine bases thymine (T) and cytosine (C) forming H-bonds (shown by dotted lines) with, respectively, the purine bases adenine (A) and guanine (G). The carbohydrate-phosphate backbone linked at carbon atom 1 of the deoxyribose to each base is on the outside of the molecule and from the X-ray data of Rosalind Franklin was shown to follow space group C2 symmetry with one chain running 5' to 3' and the other antiparallel from 5' to 3' in the opposite direction. The precise base links were worked out by James Watson and Francis Crick following Edwin Chargaff's publication in 1950 of the molar equivalence of A with T, and C with G, and Jerry O'Donohue's identification of the correct (keto rather than enol) tautomeric forms for the bases. Adapted with permission from Booth, H. and Hey, M.J. (1996) *J. Chem. Educ.* **73**, 928–931. Copyright (1996) the American Chemical Society.

Production of high-quality, non-degraded and high-purity DNA

The production of high molecular mass, undegraded DNA was the first priority, and this involved an extraction process from calf thymus that did not involve the traditional use of high acid or high alkali conditions. A mild extraction procedure using near neutral solutions and low temperature (at $\sim 0^{\circ}\text{C}$) was used, improving on the earlier procedure of Hammersten in Stockholm. The absence of protein/ degradative enzymes was checked by the well-established Biuret and Sakaguchi assays and, reassuringly, the measured N content was found to match that calculated from the base content. Macromolecular integrity was then checked by the then fledgling technique of analytical ultracentrifugation. Professor Sandy Ogston at the Department of Biochemistry in Oxford had a Svedberg analytical ultracentrifuge and offered the services of his laboratory; a sedimentation velocity experiment using Schlieren optics performed by him and his student D. Cecil (later published in 1949, also in the

Figure 2. Schlieren photograph from a sedimentation velocity experiment performed by Cecil and Ogston at Oxford on Nottingham's calf thymus DNA preparation in 0.2 M NaCl. The single hyper-sharp boundary confirms the preparation is of very high purity. Reproduced with permission of the Royal Society of Chemistry.



Journal of the Chemical Society) showed a single hyper-sharp Schlieren peak (Figure 2). A combination of the sedimentation coefficient with the diffusion coefficient gave a weight or mass average molecular mass M_w of 820,000 daltons (Da), confirming the molecule's intact status: the sample was ready for titration studies.

Titrimetric studies

It was then possible for Gulland, Jordan and Taylor to examine the titration behaviour of the amine and

Figure 3. Left: John Masson Gulland (Courtesy of Cold Spring Harbor Laboratory Library & Archives). Centre: D.O. 'Doj' Jordan. Right: PhD student J. Michael Creeth.



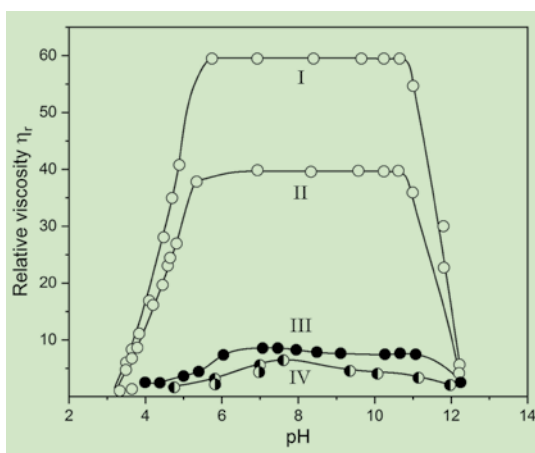
hydroxyl groups on bases of the DNA and to see if they reproduced the same behaviour of the bases in free solution. This proved not to be the case. Starting at neutral pH and adding either increasing alkali to titrate the hydroxyl groups, or increasing acid to titrate the amino groups, resulted in a marked shift of the pKa values, indicating that in the DNA molecule, the titratable groups may have been blocked by the formation of H-bonds, and only extremes of pH could disrupt this. The final and decisive experiments that followed involved hydrodynamic measurements.

Hydrodynamic studies

The final definitive measurements were performed by Creeth, Gulland and Jordan (Figure 3) using primarily capillary viscometry (using a specially designed viscometer known as a Frampton viscometer, which permitted the measurement of viscosity at different hydrostatic pressures) reinforced by streaming birefringence measurements. PhD student Michael Creeth made a succession of measurements of the relative viscosity (viscosity of a solution relative to the solvent) on the high-quality calf thymus DNA preparation (Figure 4). Starting with neutral solutions in which the relative viscosity was at a maximum, when the pH was successively increased by adding alkali, the viscosity was observed to drop dramatically above a pH of 11.9. Similarly, and again starting off at neutral pH conditions, addition of acid resulted in a dramatic viscosity reduction below pH 5.6. Both pH values corresponded to the titration data. These dramatic drops in viscosity confirmed the disruption of the DNA molecule through rupture of the H-bonds.

These observations were reinforced with the technique of streaming birefringence measurements: the alignment of linear molecules of high molecular mass causes birefringence as measured under a polarizing microscope, a birefringence that was lost at the same pH extremes as seen using viscometry and titration. This further demonstrated that H-bonds exist between nucleotides in the DNA molecule, which are disrupted at high-alkali or high-acid, as Creeth, Gulland and Jordan noted:

Figure 4. Reduction of the relative viscosity (open circles) of calf thymus DNA through the addition of acid or alkali. Starting at neutral pH, alkali was added to a pH of over 12. Then starting at neutral pH, acid was added to a pH of below 3.5. This procedure was repeated for two different hydrostatic pressures (I and II), with similar results. The dark and half circles (III and IV) correspond to DNA that had been pre-treated respectively with strong acid or alkali. Redrawn from Creeth, Gulland and Jordan's 1947 paper.



“The critical pH values are coincidental with those in which a liberation of amino and hydroxyl groups has been observed and it is considered that the two phenomena are related and due to the fission of the hydrogen bonds postulated as linking the purine-pyrimidine hydroxyl group and some of the amino groups”.

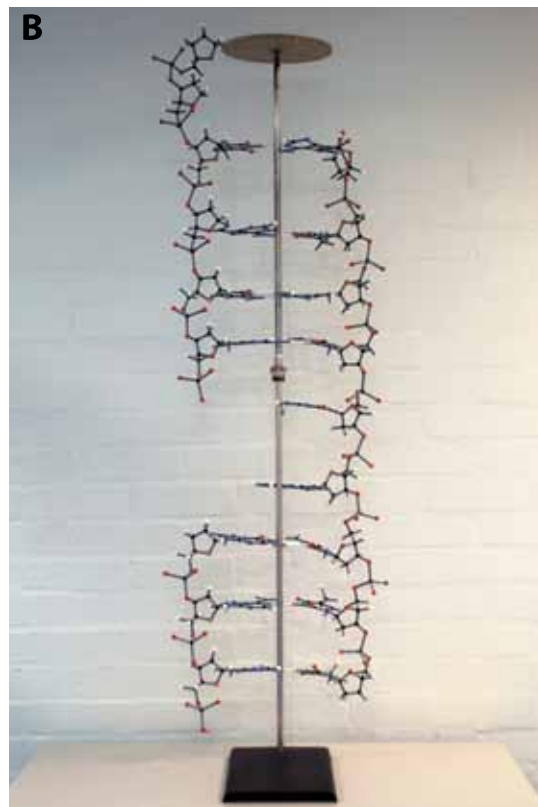
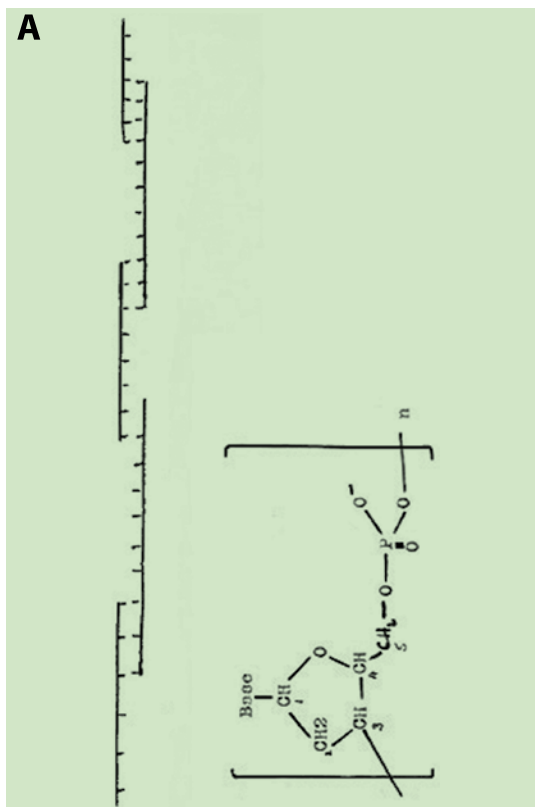
The Creeth 2-chain H-bonded model for DNA

The drop in relative viscosity together with a reduction in streaming birefringence observed upon disruption of the H-bonds could be caused by one of two changes to

the then unknown DNA structure: (i) a large reduction in molecular mass of the structure (disruption of inter-chain H-bonds linking separate chains) or (ii) a reduction of asymmetry of the structure (disruption of intra-chain H-bonds within the same chain). In his PhD thesis, which appeared shortly after the published works, Michael Creeth considered these cases and for case (i) he drew his own 2D model of what the DNA molecule might look like, including the incorporation of H-bonded bases. His drawing is reproduced in Figure 5. The model comprised (correctly) two chains, linked together by H-bonds between the bases, with the bases on the inside of the molecule and the sugar-phosphate backbones (worked out earlier by P.A. Levene in New York) correctly on the outside of the 2-chain structure.

This model—which was merely to indicate the possible arrangement of H-bonds in the DNA molecule—as it turned out, wasn't that far away from the correct model for the structure of the DNA molecule worked out by Watson and Crick six years later. Figure 5B shows a modern ball-and-stick model for Creeth's structure. The Creeth model has no double helix of course and it wasn't known which pyrimidine base linked with which purine. He had no access to X-ray data which indicated the structure was helical. Intriguingly, not long after, in 1949 Norwegian Sven Furberg, a student in J.D. Bernal's laboratory at Birkbeck College, submitted his PhD thesis which contained a helical single-chain model for DNA.

Figure 5. (A) Creeth's 2-chain model for the structure of DNA from page 85 of his PhD thesis. The chains are represented by 2 vertical lines with the bases shown as short horizontal lines on the inside, linked together by H-bonds. The backbone nucleotide unit (from page 79) is also shown. How each base linked up (A, T, C and G) was not known. **(B)** A modern ball-and-stick representation of the Creeth model in the National Centre for Macromolecular Hydrodynamics.



A recent publication has speculated what might have happened if Creeth and Furberg had met up to discuss their findings...

Creeth also included breaks with regions of unpaired H-bonds in the chains. This was presumably to explain the dramatic drop in the relative viscosity after scission of the H-bonds. A recently published study in *Biochemical Society Transactions* showed that, if Creeth had access to more modern hydrodynamic theory, the incorporation of breaks in the chains would not have been necessary to explain the observed drop in relative viscosity—a simple halving of the molecular mass with the two intact chains coming apart would have been sufficient to explain the data. It is always easier in hindsight!

Concluding remarks and postscript

The discovery of H-bonds linking the bases of the DNA molecule by the Nottingham team (Figure 3) paved the way for Watson and Crick. Once Chargaff and co-workers had published their findings in 1950 on the equivalence of the bases (A:T, C:G) and the correct tautomeric forms for the thymine and guanine (keto not enol) had been identified, Watson and Crick were able to correctly establish the H-bond links between the bases (Figure 1) and complete their model for the structure of DNA.

There is a nice postscript. Michael Creeth's earlier model with its alternating breaks along the two chains bears close similarity to modern textbook representations of how, during the replication cycles of some viruses such as the bacteriophage Mu, there are staggered cuts in the DNA duplex with resulting short single-stranded segments; and in the generation of double-stranded DNA from retroviral single-stranded RNA. It also bears similarity to drawings that could depict the design of polymerase chain reactions with multiple annealing primers. ■



Stephen Harding is Professor of Applied Biochemistry at the University of Nottingham and has been Director of the National Centre for Macromolecular Hydrodynamics (NCMH) since its foundation in 1987. He develops and applies hydrodynamic methodology

to the understanding of the sizes and shapes of biological macromolecules. He is also Adjunct Professor at the University of Oslo where he is helping design bioinspired polymers for preserving the Oseberg Viking ship and its artefacts. He was trained by Colin Blake at the University of Oxford, Arthur Rowe at the University of Leicester, Paley Johnson at the University of Cambridge and Michael Creeth at the University of Bristol. He was Creeth's last postdoc before his retirement.

Email: Steve.Harding@nottingham.ac.uk

Further reading

- Harding, S.E., Channell, G. and Phillips-Jones, M.K. (2018) The discovery of hydrogen bonds in DNA and a re-evaluation of the 1948 Creeth two-chain model for its structure. *Biochem. Soc. Trans.* **46**, 1171–1182
- Harding, S.E. (2019) DNA Day 2019: Nottingham's discovery of hydrogen bonds in DNA. YouTube: <https://www.youtube.com/watch?v=wQHYSixvKVY> and <https://www.youtube.com/watch?v=1Ms77LkQBrY>
- Watson, J.D. (2012) *The Annotated and Illustrated Double Helix* (Gann, A. & Witkowski, J., eds), Simon & Schuster, New York, USA
- Creeth J.M. (1948) Some physico-chemical studies on nucleic acids and related substances. PhD Dissertation, University of London <http://ethos.bl.uk/OrderDetails.do?did=1&uin=uk.bl.ethos.729338>
- Gulland, J.M., Jordan, D.O. and Threlfall, C. (1947) Deoxypentose nucleic acids, Part I. Preparation of the tetrasodium salt of the deoxypentose nucleic acid of calf thymus. *J. Chem. Soc.*, **212**, 1129–1130
- Gulland, J.M., Jordan, D.O. and Taylor, H.F.W. (1947) Deoxypentose nucleic acids, Part II. Electrometric titration of the acidic and basic groups of the deoxypentose nucleic acid of calf thymus. *J. Chem. Soc.* **213**, 1131–1141
- Creeth, J.M., Gulland, J.M. and Jordan D.O. (1947) Deoxypentose nucleic acids, Part III. Viscosity and streaming birefringence of solutions of the sodium salt of the deoxypentose nucleic acid of calf thymus. *J. Chem. Soc.* **214**, 1141–1145
- Cecil, D. and Ogston, A.G. (1948) The sedimentation of thymus nucleic acid in the ultracentrifuge. *J. Chem. Soc.* **280**, 1382–1386
- Harding, S.E. (1997) The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution. *Prog. Biophys. Mol. Biol.* **68**, 207–262
- Madigan M.T. and Martinko J.M. (2006) *Viral Diversity*, Brock Biology of Microorganisms, Pearson Prentice Hall, New Jersey, USA
- Williams, G. (2019) *Unravelling the Double Helix*, Weidenfeld & Nicholson, London, UK