Fifty Years Ago

DNA: the double helix

Giorgio Semenza[a,b]

[a] *Dipartimento di Chimica, Biochimica e Tecnologie per la Medicina, Università di Milano, Via Saldini 50, I-2013 Milano, Italy*
[b] *Swiss Institute of Technology, Biochemistry, ETH-Zentrum, POB 35, CH-8092 Zürich, Switzerland*

'The chemical nature of the genes is unknown. They are probably proteins, because the nucleic acids consist of (few) blocks of tetranucleotides (Levene, 1930). This simple structure makes them unlikely candidates as carriers of the large information needed to act as genes.' This view –strange as it may seem today – was shared by most biologists when I was preparing my exam in biology at the medical school in 1946.

In a series of classical papers from 1906 on, Levene's group and others had identified D-ribose and 2-deoxy-D-ribose as the sugars occurring in RNA and in DNA, respectively. They had also identified the purines and pyrimidines, localized the phosphodiester link between the sugars, and classified the linkage with the bases as glycosidic. Levene's work culminated in the suggestion of a tetranucleotide structure for both RNA and DNA, i.e. that the nucleic acids are composed of an (unknown) number of building blocks, the tetranucleotides, each consisting of four sugars, four phosphates and one each of the four bases. In his words, 'the tetranucleotide theory is the minimum molecular weight and the nucleic acid may as well be a multiple of it' (Levene and Bass, 1931, p. 289 [8]). Curiously, the analytical data did not indicate equal ratios among the four bases (e.g. 1:0.6:1.2:0.8); the deviation from 1:1:1:1 ratios was attributed to the lack of precision of the precipitation procedures used to isolate the individual bases. The treacherous 'beautiful simplicity' of the tetranucleotide structure and Levene's authority effectively stopped the search for alternative structures for the nucleic acids and, what is worse, proved to be an additional hurdle to the (now 'classical') work of Avery, MacLeod and McCarty (1944) [1] to rapidly and universally be accepted.

As most readers know, Avery et al. purified a very viscous substance from *Diplococcus pneumoniae*, which transformed a strain of bacteria into another with immunologically different polysaccharide capsulae. The transformed strain was genetically stable. The highly viscous transforming factor, purified by a very accurate and mild procedure, had the characteristics of DNA: aside from the evidence from chemical analysis, it was not affected by ribonuclease, trypsin or chymotrypsin, but was inactivated by (crude) DNase. (Four years later, this was confirmed with use of crystallized DNase made available to them by Kunitz.) Avery et al.'s results were soon confirmed by other groups and extended to transformations in other bacteria. Leading geneticists had no difficulty in accepting the idea that these transformations were due to the chemical isolation and transfer of genes, and not to the induction of a mutation. Avery et al.'s observations were indeed a boost to bacterial genetics, a field then in its infancy.

Still, many – among them the authoritative Mirsky – for a long time regarded Avery et al.'s evidence as insufficient: the transforming activity could still be attributed to a very minor component in the DNA preparation. However, by 1952 active preparations were obtained in which the protein content was less than 0.02%. This, plus the work of Luria, Delbrück, and others, on bacteriophages, especially the Hershey and Chase experiment showing that bacteriophages injected their DNA but not the protein coat into the bacteria they infected, gave the final evidence that genes are indeed made of DNA. This conclusion, therefore, threw serious doubts on the correctness of Levene's tetranucleotide structure of DNA.

In the late forties Chargaff – working a few miles north of Avery's lab – correctly identified the weakest point in the tetranucleotide structure: do the four bases really occur in 1:1:1:1 ratios in DNA and in RNA? To settle this decisive point, Chargaff's group worked out what had been lacking until then: a quantitative procedure to measure the purine and pyrimidine bases. Mild preparation procedures (comparable to those of F. Miescher, E. Hammarsten, R. Signer, and others), leading to highly viscous DNA; reliable acid hydrolysis, paper chromatography; elution of the ultraviolet light (UV)-absorbing spots and their spectrophotometric determination made possible reliable, quantitative determinations of the base composition of RNAs and DNAs. By 1950 and 1951 Chargaff [2,3] could summarize extensive analytical data on the base composition of DNAs from a number of organs (e.g. thymus, liver, spleen), from human sperm, from yeast, and from a variety of bacteria. In most DNAs the four 'classical' bases (adenine, guanine, cytosine and thymine) occurred in ratios clearly different from 1:1:1:1 (Chargaff, 1950 [2]). That was the end of Levene's tetranucleotide structure.

Naturally, Chargaff did not fail to realize the importance of the newly discovered ratios among the purine and pyrimidine bases in the DNAs: the A:T and G:C ratios were never significantly different from 1, whereas the A:G and the T:C varied widely from DNA to DNA. Shortly thereafter Pauling and Corey (1953) [11] suggested that the DNA is composed of three polynucleotide chains coiled and intertwined into a helix, with the phosphoryl groups closely packed about the axis of the column, and the nitrogen bases projecting radially.

One can wonder why Chargaff did not himself suggest a structure for DNA incorporating in some way the remarkably constant A:T and G:C ratios of 1 to 1. No doubt, he resented that 'strangers' such as Watson and Crick were coming up with a structure (1953) [15], which dangerously showed from the beginning a likely chance of being right. I happen to have been present at some unusual and remarkable verbal clashes ('Real research is done at the bench, not by playing about with metal models') – Chargaff's in-depth command of the English language and its sharp use came to the fore all too often. But, in all fairness, since he had himself witnessed how paralyzing a wrong structure (Levene's) had been for the development of the field during nearly two decades, one can

E-mail address: semenza@bc.ciol.ethz.ch (G. Semenza).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00471-X

perhaps understand that Chargaff disliked the idea of any structure not arising 'spontaneously', almost per se, from the raw data themselves.

There is no need for me to repeat here what has been written on how Watson and Crick came to 'A structure for deoxyribose nucleic acid' (the title of their one-page paper, 1953 [15]) – which was followed back to back by the two papers from M.H.F. Wilkins et al. (1953) [17], and R.E. Franklin and G.R. Gosling (1953) [5]. Few bioscientists have not read Watson's 'The Double Helix' (1968) [14], its sequelae (see e.g. the latest by B. Maddox, 2002 [9]), and what has been written or reported in a number of symposia, meetings, even daily papers, but most outstandingly by Judson (1996) [7].

There can be no doubt that Watson and Crick's papers of April of 1953 [15], and perhaps even more so that of May (1953) [16] stand out in the bioscience of the 20th century as much as Darwin's 'Origin of Species' stood out in the bioscience of the 19th century.

Like Darwin's book, Watson and Crick's papers produced a major turn in biological thinking. Darwin did away with the literal interpretation of the Bible insofar as it dealt with appearance and change of life on the earth, and also with incorrect attempts to explain these changes. When Watson and Crick presented their DNA double helix, Levene's tetranucleotide had been buried for some years, but Pauling and Corey's tri-helical structure was just being proposed (1953) [11]. Curiously, the comparison between Darwin's and Watson and Crick's publications extends to some of the circumstances under which their works were published. Without the nearly identical ideas of a friendly competitor (Wallace) and the encouragement of authoritative scientists, Darwin would have probably hesitated a few more years before publishing the *Origin of Species*. In Watson and Crick's situation, Wallace's role was played by Wilkins, Franklin, and their associates.

Naturally, in performances of this magnitude some details may be found to be wanting – but it is the open vistas which count. Sure, Watson and Crick's original structure (1953) [15] 'was built' with only two hydrogen bonds between cytosine and guanine (the third bond would be suggested by Pauling and Corey three years later, in 1956 [12]). Also, Watson and Crick did not consider the minor methylated or hydroxymethylated bases (which would be discovered shortly after), but the extra group(s) do not interfere with the Watson–Crick base pairing.

To Darwin's strict survival of the fittest the following decades added additional concepts (see, e.g. [4,6] and others), but the validity of le Hasard et la Nécéssité has never been questioned.

It was clear to Watson and Crick themselves that their DNA double helix was opening a new era. In the April paper (1953) [15] one of the last paragraphs is a classic example of scientific understatement: 'It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material'. This is the subject of a paper which Watson and Crick published a month later (1953) [16]; this second paper presented a number of 'speculations', all of which subsequent years proved to be remarkably true.

'In a long molecule (of DNA) many different permutations are possible, and it seems therefore likely that the precise sequence of the bases is the code which carries the genetic information. If the actual order of the bases of one of the pair of chains were given, one could write down the exact order of the bases in the other one, because of the specific pairing. Thus one chain is, as it were, the complement of the other, and it is this feature which suggests how the deoxyribonucleic acid molecule might reproduce itself...' The DNA is 'in fact, a pair of templates, each of which is complementary to the other'. (Meselson and Stahl would elegantly show in 1958 this to be actually the case [10].) 'We imagine that prior to duplication the hydrogen bonds are broken, and the two chains unwind and separate.' The double helix is indeed a beautiful example of how astute nature can be: the two polynucleotide chains are kept together by the hydrogen bonds between the base pairs, while the negative charges of the phosphates would tend to pull them apart.

'Our model suggests... that spontaneous mutation may be due to a base occasionally occurring in one of its less likely tautomeric forms' – again a suggestion which is now an established fact.

'Our structure... is an open one. There is room between the pair of polynucleotide chains for a polypeptide chain to wind around the same helical axis. It may be significant that the distance between adjacent phosphorous atoms, 7.1 Å, is close to the repeat of a fully extended polypeptide chain.' Watson and Crick stopped short of suggesting that proteins can interact with the bases in intact double stranded DNA – as repressors, transcription factors, or for DNA methylation, etc.

I have taken the liberty of quoting *verbatim* from Watson and Crick's prose because 50-year-old papers are seldom available in our overcrowded libraries. More importantly, their two 1953 papers have indicated at a very early stage which direction molecular biology and biochemistry would follow for a number of decades.

I am told that the day they had finally worked out the details of the DNA double helical structure, Watson and Crick went to a pub near the Cavendish laboratory and announced: 'Today we have discovered the secret of life!' I think that they were not far from the truth. But the present pub keeper is remarkably ignorant of this incident and is innocent of basic molecular biology: 'Watson and Crick, you said? Never heard of them' (Tanford, 2003) [13].

PS. However, on the occasion of this year's festivities for the DNA double helix a plaque has been uncovered at the Eagle Pub in the presence of Jim Watson.

Acknowledgements: I wish to thank the feedback and help from my somewhat younger colleague Ned Mantei, for whom Watson and Crick ranked high in his list of 'scientific heroes' during his student years in the 1960's.

References

[1] Avery, O.T., MacLeod, C.M. and McCarty, M. (1944) J. Exp. Med. 79, 137–158.
[2] Chargaff, E. (1950) Experientia 6, 201–209.
[3] Chargaff, E. (1951) J. Cell. Comp. Physiol. 38 (Suppl. 1), 41–59.
[4] Dawkins, R. (1976) The Selfish Gene, Oxford University Press, New York.
[5] Franklin, R.E. and Gosling, R.G. (1953) Nature 171, 740–741.
[6] Gould, S.J. (2002) The Structure of Evolutionary Theory, Belknap Press of Harvard University Press, Cambridge, MA.
[7] Judson, H.F. (1996) The Eighth Day of Creation (expanded edi-

tion), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[8] Levene, P.A. and Bass, L.W. (1931) Nucleic Acids, Chemical Catalogue Co, New York.
[9] Maddox, B. (2002) The Dark Lady of DNA, Harper and Collins, New York.
[10] Meselson, M. and Stahl, F.W. (1958) Proc. Natl. Acad. Sci. USA 44, 671–682.
[11] Pauling, L. and Corey, R.B. (1953) Proc. Natl. Acad. Sci. USA 39, 84–97.
[12] Pauling, L. and Corey, R.B. (1956) Arch. Biochem. Biophys. 65, 164–181.
[13] Tanford, Ch. (2003) in: Personal Recollections, VII; Compr. Biochem. Vol. 42 (Semenza, G. and Turner, A.J., Eds.), pp. 1–52, Elsevier Science Publishers, Amsterdam.
[14] Watson, J.D. (1968) The Double Helix, Atheneum, New York.
[15] Watson, J.D. and Crick, F.H.C. (1953) Nature 171, 737–738.
[16] Watson, J.D. and Crick, F.H.C. (1953) Nature 171, 964–967.
[17] Wilkins, M.H.F., Stokes, A.R. and Wilson, H.R. (1953) Nature 171, 738–740.