Series

What history tells us XIII. Fifty years of the Central Dogma

Michel Morange

Centre Cavaillès and IHPST, Ecole normale supérieure, 29 rue d'Ulm, 75230 Paris Cedex 05, France

(Fax, 33-144-323941; Email, morange@biologie.ens.fr)

1. Introduction

It is now 50 years since the publication of a lecture held one year before in which Francis Crick put forward two concepts which, together with Darwin's principle of natural selection, are believed to provide the underpinning to all of biology: the Central Dogma and the Sequence Hypothesis (Crick 1958). This event has been commemorated by historians and philosophers of science (Fantini 2006). The Central Dogma occupies a pre-eminent place in molecular biology, and other disciplines have subsequently also defined their own Central Dogmas: neurobiology (Colucci-D'Amato *et al* 2006) and cell biology (Cooper 1981); and even flow cytometry (Ornstein 1983) and palliative care (Gillick 2005)! But the Central Dogma has also been harshly criticised by opponents of molecular biology.

The 50th anniversary offers an excellent opportunity to reconsider the Central Dogma, what it did and did not assert, what exactly Crick said, the context in which it was proposed, the challenges it has faced and overcome or failed to overcome since its inception, and the roots of its legitimacy.

2. The Central Dogma of molecular biology

Crick worded the Central Dogma thus: "This states that once 'information' has passed into protein *it cannot get out again*. In more detail, the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information means here the *precise* determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein" (Crick 1958). To give a full account of the complex relations existing between macromolecules present in organisms, it had to be complemented by the Sequence Hypothesis, which went "In its simplest form it assumes that the specificity of a piece of nucleic acid is expressed solely by the sequence of its bases, and that this sequence is a (simple) code for the amino acid sequence of a particular protein" (Crick 1958) and by the mechanism of protein folding that Crick proposed a few pages before: "…the more likely hypothesis is that the *folding is simply a function of the order of the amino acids*, provided it takes place as the newly formed chain comes off the template" (Crick 1958).

Two points deserve discussion: the choice of the word 'dogma', and the meaning of 'information'. Concerning 'dogma', Jacques Monod was the first to tell Crick that "a dogma is something which a true believer cannot doubt" (Judson 1996), probably not what Crick meant. On different occasions Crick replied that, from his distrust in religious beliefs in general, by dogma he only meant a bold hypothesis with no experimental support (Crick 1988; Judson 1996). During his whole life, Crick emphasized the importance of theories and models in guiding experimental work and helping to eliminate lines of research leading to dead ends. The choice of the word dogma was probably not ideal, but the notion of information was, in contrast, very clearly defined: a precise determination of sequence. Therefore what the Central Dogma said exactly was that a sequence of nucleotides can determine a sequence of amino acids, but that a sequence of amino acids cannot determine a sequence of nucleotides. The reason in favour of the latter exclusion was the apparent total absence of machinery able to translate the sequence of a protein into the sequence of a nucleic

Keywords. Chaperone; Central Dogma; evolution; prion; reverse transcriptase

acid, combined with the extreme unlikelihood that the existing (very likely complex) machinery, which transferred information from nucleic acid to protein, could be run in reverse. In addition, protein folding made the sequence of a protein inaccessible to any obvious means by which the sequence might be copied. The fact that it was impossible to transfer information (with the meaning given by Crick to this term) from protein to nucleic acid turned out to be the molecular equivalent of something else that was also impossible, namely for the phenotype to specifically alter the genotype, or for the soma to modify the germ line. This was not noted immediately but only later by commentators including Monod (Judson 1996): the Central Dogma fitted perfectly well the Weismannian neo-Darwinian view.

3. The context in which the Central Dogma was enunciated

The main objective of the Central Dogma was to introduce a simple order into the complex relations between macromolecules that had emerged from numerous observations made over the previous years. The richness, but also the complexity, of the relations proposed by Jean Brachet (Thieffry and Burian 1996), for example, is in sharp contrast with the objectives of Crick. His efforts were not yet totally successful, even in his own mind. One sees this from the interpretation drawn by him from many experiments, that "protein and RNA require common intermediates for their synthesis" (Crick 1958; it should be pointed out that this was before the distinction between messenger, ribosomal and transfer RNA had acquired the clarity that it did later). There is no doubt, though, that the assumptions made by Crick in formulating the Central Dogma were of enormous heuristic value in working out many of the details of information transfer from genes to proteins, though not in the elucidation of the actual genetic code. Crick found support for his efforts in the reconstitution experiments recently done by Heinz Fraenkel-Conrat showing that the RNA component of tobacco mosaic virus was alone able to determine the specific characteristics of the viral coat protein, but that the reverse was not true, and in the results obtained by Christian Anfinsen on the spontaneous refolding of ribonuclease (Strasser 2006). Soon the Central Dogma became fashionable, and in the process lost its precise meaning. The version publicised by James Watson in 1965, with one curved arrow beginning and ending at DNA (signifying duplication), and two unidirectional arrows, one each running from DNA to RNA and RNA to protein, replaced the cautious wording of Crick (in Watson's text, the phrasing was '…RNA sequences are never copied on protein templates; likewise, RNA never acts as a template for DNA'; Watson 1965).

Two other meanings of the word 'information' were added to the one initially proposed by Crick. Information was said to include the complexity of the 3-dimensional structure of proteins as well as the regulatory processes occurring in organisms. The Central Dogma was progressively transformed into a "Central Dogma of Life" (Schreiber 2005), thereby becoming essentially a metaphor for the reductionist approach characteristic of molecular biology.

4. Challenges to the Central Dogma

I will successively examine four sets of discoveries and observations that have been considered as challenges to the Central Dogma: the discovery of reverse transcriptase, the role of prions in spongiform encephalopathies, the role of chaperones in protein folding, and new phenomena occurring in the transfer of information from DNA to proteins (epigenetic modifications of DNA, RNA interference, RNA splicing and RNA editing).

The first discovery was of an enzyme that could catalyse the synthesis of DNA using RNA as a template (reverse transcriptase), made by Howard Temin and Satoshi Mizutani and, independently, by David Baltimore (Temin and Mizutani 1970; Baltimore 1970). Already in their 1970 publication Temin and Mizutani drew attention to the possibility that it could have 'strong implications…for theories of information transfer in other biological systems'. An anonymous commentator in *Nature* thought it contradicted the Central Dogma (Anonymous 1970), and Temin said so too five years later in his Nobel lecture (Temin 1976). However, as argued by Crick (Crick 1970) and shown subsequently by Lindley Darden (Darden 1995), it is obvious that the existence of a reverse transcriptase did not go against the Central Dogma as had been formulated by Crick, but only against the version popularised by Watson. When Temin proposed (initially as early as 1964) that RNA viruses copied their RNA genome into DNA before integrating it into the genome of the host (whence the name 'retrovirus'), the opposition to his proposal had not been because of a blind belief in the Central Dogma. Rather, it was based on the absence of clear experimental support in its favour (Marcum 2002) – which also explains its rapid acceptance when reverse transcriptase was discovered. In fact in his original formulation of the Central Dogma, Crick (1958) had not excluded the possibility of a transfer of information from RNA to DNA. But he considered it a rare event. It was not the same for Temin, who thought that a reverse transcriptase was present in every cell, and that the genome was permanently remodelled by insertion of the DNA copies of the most actively expressed genes (Temin 1971). Today the genomes of higher organisms are acknowledged to contain an abundance of 'selfish' DNA sequences, many of which appear to have originated via reverse transcription from intruding viruses. Still, the heretical hypothesis of

Temin – of a significant modification of the genome – has not so far found experimental support.

The discovery of protein-only pathogenic agents has also been considered as a blow to the Central Dogma. As in the previous example, it is easy to show that it was not the Central Dogma enunciated by Crick that was challenged, but a more popular version in which the word "information" had lost its precise meaning of "determining sequence" and simply stood for the complex 3D structure of proteins (Keyes 1999; Hunter 1999). As early as 1967, J S Griffith proposed different models able to explain the replicative power of a protein-only infectious agent fully compatible with the principles of molecular biology; indeed one of them is very similar to the presently accepted model (Griffith 1967). In his 1970 article, by alluding to "…the problem of the chemical nature of the agent of the disease scrapie", Crick showed that he was aware of the need for examining the implications of the existence of infectious agents devoid of nucleic acid for the Central Dogma (Crick 1970). An interesting parallel with the episode of reverse transcriptase is that Stanley Prusiner was largely responsible for the idea that his discovery of prions challenged the Central Dogma of molecular biology. Not only because he wrote it, but because the first models he proposed to explain how a protein could be a pathogenic agent did challenge the Central Dogma of molecular biology: in particular, he hypothesised that a protein could generate a perfect copy of itself, with an identical sequence (Prusiner 1982, 1984; Morange 2007). That hypothesis vanished, but, falsely, the feeling that the Central Dogma had been violated remained.

Similarly, the discovery of chaperones did not reverse the Central Dogma (Morange 2005). Superficially, it appeared to falsify the hypothesis made by Crick in his 1958 publication, that "folding is simply a function of the order of the amino acids". It did not, because it was rapidly shown that in a formal sense, in most cases molecular chaperones do not possess steric information. They only prevent the formation of non-functional 3D protein structures (Ellis and Hemmingsen 1989). It is nevertheless obvious that the discovery of chaperone function was delayed by the conviction that, since (some) proteins *can* fold *in vitro* in the absence of any other protein, all proteins *must* fold within cells without any assistance.

A stronger challenge to the Central Dogma apparently came from the numerous observations made since the 1970s, which obscured the simple rules of information transfer described at the end of the 1950s. This had to do with the discovery of splicing (the protein sequence cannot be read off directly from the primary DNA sequence), editing (the sequence of RNA is altered after its copy on DNA by modification or insertion of bases), DNA methylation, and the major role of small RNAs in the regulation of gene expression. Once again, it is necessary to distinguish between what was in the Central Dogma as originally stated by Crick, and what was in later formulations by others, generally with the aim of encompassing it within a vague molecular paradigm. The confusion in this field is extreme. The evidence for DNA methylation and the discovery of microRNAs and of their regulatory functions do not really challenge the Central Dogma as expressed by Crick. The editing and splicing processes are more interesting – and challenging (Jukes 1990; Thieffry and Sarkar 1998). Proteins (protein complexes used in editing and trans-splicing factors) can alter the sequence of messenger RNAs. Is it sufficient to consider that Crick's version of the Central Dogma has been reversed (Stotz 2006)? Opposite views have been expressed (Rosenberg 2006). Is the precise sequence of amino acids converted into a sequence of RNA during the editing and splicing processes? The answer is clearly no. What is observed is that in certain relatively rare cases, proteins can affect the sequence of nucleic acids (RNA molecules), select between different possible sequences, or transform one sequence into another. In a formal sense, this is no more a 'specification of a nucleic acid sequence by the amino acid sequence in a protein' than is the phenomenon of DNA proof-reading and repair, during which (protein) enzymes make sure that 'erroneous' bases are removed from DNA.

5. The Central Dogma in an evolutionary perspective

The vision of the Central Dogma proposed by Francis Crick has resisted the many different challenges it has faced, and the numerous phenomena discovered since its inception. Paradoxically, the reasons behind its traditional justifications have proven fragile. True, no mechanism able to convert a protein sequence into a nucleic acid sequence has been discovered so far. But one would do well to remember that another mechanism, that of protein folding, remained invisible for more than thirty years before it began to occupy a central place in the descriptions of cell biologists. The Weismannian justification of the Central Dogma is also tenuous: after all, the functional separation between the soma and the germ line is strictly true in only a small part of the living world. In any case, trying to legitimise the Central Dogma by appealing to the principle that the phenotype does not modify the genotype is not satisfying.

I should like to suggest that the only possible justification for Crick's version of the Central Dogma is a justification that has not yet been seriously considered. It is based on the evolutionary relation (both inferred and surmised) that exists between DNA, RNA and proteins. The idea that nucleic acids, probably RNAs, preceded proteins in prebiotic times emerged early (Woese 1967; Crick 1968; Orgel 1968). But one had to wait for the demonstration that RNAs are able to catalyse chemical reactions as

efficiently as proteins for the hypothesis of an RNA world to emerge, and a strongly argued evolutionary history of the different classes of macromolecules to be constructed. The discovery that the peptide bond is formed by the RNA component of the ribosome was a complete surprise, and it provided strong support for the hypothesis that proteins were progressively constructed by organisms belonging to an RNA world. The extremely sophisticated nature of the biochemical reactions needed for the synthesis of deoxyribonucleotides from nucleotides is apparently out of reach of the most complex ribozymes. If true, it implies that the replacement of RNA by DNA as a genetic material followed the replacement of ribozymes by enzymes formed of amino acids.

This is a widely accepted picture today, even if other pictures persist. What was the raison d'être behind the changes in the roles played by these macromolecular actors in the drama of heredity and information transfer? Thanks to the diversity of chemical functions that amino acids and peptides can perform, the elaboration of proteins permitted an extension of the functional capacities of the first cells. On the other hand, and for at least two different and complementary reasons, the appearance of DNA generated a more stable bearer of genetic information: deoxyribose is chemically less reactive than ribose, and the double helical structure of DNA permits an inbuilt proofreading process that single RNA chains do not possess (Lazcano *et al* 1988). Patrick Forterre (Forterre 2006) has proposed that the first, RNA-based, organisms were equipped with mechanisms to protect themselves against 'foreign RNA' (the function of the mechanisms would have been similar to the function of restriction enzymes in today's DNA world). By converting their genetic material to DNA, viruses would have circumvented those mechanisms. In a sense, therefore, viruses may have initiated the process that led to the replacement of the RNA world by a DNA world. The higher stability and efficient proofreading of DNA may have been later adaptations. Whatever the actual sequence of events, it is clear that the concomitant existence of passages from RNA to DNA and from DNA to RNA is necessary if the presently accepted picture of early evolution is correct. Conversely, the construction of proteins with information contained in RNA did not require an opposite movement of information from proteins to RNA. It did not exclude it; but in a relatively stable environment, there would have been no evolutionary pressure in its favour: all that would be required is that the same proteins be reproduced at each generation through a reliable genetic code. The existence of editing and trans-splicing enzymes and proteins shows that it was not a "principle" that prevented the transfer of sequence information from proteins to nucleic acids, but rather their evolutionary history, which moulded the relations between the three classes of macromolecules.

6. Conclusion

For Carl Woese, the "failure to embrace evolution is the dark side or Achilles' heel of molecular biology"; "Molecular biology has to bring evolution to the fore and to integrate it fully – not to hold it at arm's length" (Woese 2001). Here Woese is referring to our attempts to understand translation, and doing so is pursuing an avenue he opened more than thirty years ago. But his exhortation is even more pertinent in the context of evaluating the significance of the Central Dogma in molecular biology. The version of the Central Dogma proposed by Crick has not been seriously challenged. But its justification cannot be found on any mechanistic grounds or by appealing to "principles", but only by considering the evolutionary history that has shaped the relations between DNA, RNA and proteins (Morange 2006).

Acknowledgements

I am indebted to David Marsh for critical reading of the manuscript.

References

Anonymous 1970 Central dogma reversed; *Nature (London)* **226** 1198–1199

Baltimore D 1970 RNA-dependent DNA polymerase in virions of RNA tumor viruses; *Nature (London)* **226** 1209–1211

Colucci-D'amato L, Bonavita V and di Porzio U 2006 The end of the central dogma of neurobiology: stem cells and neurogenesis in adult CNS; *Neurol. Sci.* **27** 266–270

Cooper S 1981 The central dogma of cell biology; *Cell Biol. Int. Rep.* **5** 539–549

Crick F H C 1958 On protein synthesis; *Symp. Soc. Exp. Biol.* **12** 138–163

Crick F H C 1968 The origin of the genetic code; *J. Mol. Biol.* **38** 367–379

Crick F H C 1970 Central dogma of molecular biology; *Nature (London)* **227** 561–563

Crick F H C 1988 *What mad pursuit* (New York: Basic Books)

Darden L 1995 Exemplars, abstractions, and anomalies. Representations and theory change in Mendelian and molecular genetics; in *Concepts, theories and rationality in the biological science* (eds) G Wolters and J G Lennox (Pittsburg: University of Pittsburg Press) pp 137–158

Ellis R J and Hemmingsen S M 1989 Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures; *Trends Biochem. Sci.* **14** 339–342

Fantini B ed. 2006 History of Central Dogma of molecular biology and its epistemological status today, Geneva, February 22–23, 2007; *Hist. Philos. Life Sci.* **28** 487–609

Forterre P 2006 The origin of viruses and their possible roles in major evolutionary transitions; *Virus Res.* **117** 5–16

Gillick M R 2005 Rethinking the central dogma of palliative care; *J. Palliat. Med.* **8** 909–913

Griffith J S 1967 Self-replication and scrapie; *Nature (London)* **215** 1043–1044

Hunter N 1999 Prion diseases and the central dogma of molecular biology; *Trends Microbiol.* **7** 265–266

Judson H F 1996 (1979) *The eighth day of creation. The makers of the revolution in biology* (New York: Cold Spring Harbor Laboratory Press)

Jukes T H 1990 Quakes hit California and, perhaps, the central dogma; *J. Mol. Evol.* **30** 1–2

Keyes M E 1999 The prion challenge to the "Central Dogma" of molecular biology, 1965–1991; *Stud. Hist. Philos. Biol. Sci.* **30** 1–19 and 181–218

Lazcano A, Guerrero R, Margulis L and Oro J 1988 The evolutionary transition from RNA to DNA in early cells; *J. Mol. Evol.* **27** 283–290

Marcum J A 2002 From heresy to dogma in accounts of opposition to Howard Temin's DNA provirus hypothesis; *Hist. Philos. Life Sci.* **24** 165–192

Morange M 2005 The discovery of chaperone function; *J. Biosci.* **30** 461–464

Morange M 2006 The protein side of the central dogma: permanence and change; *Hist. Phil. Life Sci.* **28** 513–524

Morange M 2007 The progressive construction of a mechanism for prion diseases; *J. Biosci.* **32** 223–227

Orgel L E 1968 Evolution of the genetic apparatus; *J. Mol. Biol.* **38** 381–393

Ornstein L 1983 The "central dogma" of flow cytometry; *Cytometry* **3** 459–460

Prusiner S B 1982 Novel proteinaceous infectious particles cause scrapie; *Science* **216** 136–143

Prusiner S B 1984 Prions; *Sci. Am.* **251** 48–57

Rosenberg A 2006 Is epigenetic inheritance a counter example to the central dogma?; *Hist. Philos. Life Sci.* **28** 549–566

Schreiber S L 2005 Small molecules: The missing link in the central dogma; *Nature Chem. Biol.* **1** 64–66

Stotz K 2006 Molecular epigenesis: Distributed specificity as a break in the central dogma; *Hist. Philos. Life Sci.* **28** 533–548

Strasser B J 2006 A world in one dimension: Linus Pauling, Francis Crick and the central dogma of molecular biology; *Hist. Phil. Life Sci.* **28** 491–512

Temin H M 1971 The protovirus hypothesis; *J. Natl. Cancer Instit.* **46** III–VIII

Temin H M 1976 The DNA provirus hypothesis: The establishment and implications of RNA-directed DNA synthesis; *Science* **192** 1075–1080

Temin H M and Mizutani S 1970 RNA-dependent DNA polymerase in virions of Rous sarcoma virus; *Nature (London)* **226** 1211–1213

Thieffry D and Burian R M 1996 Jean Brachet's alternative scheme for protein synthesis; *Trends Biochem. Sci.* **21** 114–117

Thieffry D and Sarkar S 1998 Forty years under the central dogma; *Trends Biochem. Sci.* **23** 312–316

Watson J D 1965 *Molecular biology of the gene* (New York: W A Benjamin)

Woese C 1967 *The genetic code: The molecular basis for genetic expression* (New York: Harper and Row)

Woese C R 2001 Translation: In retrospect and prospect; *RNA* **7** 1055–1067

*e*Publication: 14 May 2008