Hist. Phil. Life Sci., 28 (2006), 491-512

A World in One Dimension: Linus Pauling, Francis Crick and the Central Dogma of Molecular Biology

Bruno J. Strasser

Program in the History of Science and Medicine
Yale University
New Haven, CT, 06520-8015, USA

ABSTRACT – In 1957, Francis Crick outlined a startling vision of life in which the great diversity of forms and shapes of macromolecules was encoded in the one-dimensional sequence of nucleic acids. This paper situates Crick's new vision in the debates of the 1950s about protein synthesis and gene action. After exploring the reception of Crick's ideas, it shows how they differed radically from a different model of protein synthesis which enjoyed wide currency in that decade. In this alternative model, advocated by Linus Pauling and other luminaries, three-dimensional templates directed the folding of proteins. Even though it was always considered somewhat speculative, this theory was supported by a number of empirical results originating in different experimental systems. It was eventually replaced by a model in which the forms and shapes of macromolecules resulted solely from their amino acid sequence, dramatically simplifying the problem of protein synthesis which Crick was attempting to solve in 1957.

KEYWORDS – Central Dogma, Molecular Biology, Protein Synthesis, Protein Folding, Template, Sequence

Introduction

The world comes in three dimensions. Yet, the living world is perpetuated in one. The bewildering diversity of forms and shapes that constitutes life, from macromolecules to whole organisms, is specified in a linear sequence of simple units. This perplexing vision of life is one of the key landmarks in the intellectual history of the life sciences in the twentieth century. Today, the success of genomics and bioinformatics testifies to the power of this original conception of nature. In 1957, Francis Crick, one of molecular biology's leading theorizers, played an essential role in promoting this new vision by formulating several bold propositions, including one he called 'the Central dogma'. He defined the problem of protein synthesis, or how cells produce a wide variety of molecular structures, as a question about the transfer of one-dimen-

sional information from nucleic acids to proteins. In this sense, he departed from previous conceptions, such as Linus Pauling's, where three-dimensional structures served as molds or templates for the production of other three-dimensional structures. This shift represented a major step in the reductionist agenda of the life sciences, since it reduced some of the major problems of biology to a single dimension (Morange 2000, chapter 13).

In exploring how this particular change came about and attempting to situate the central dogma in the intellectual history of the life sciences, I will argue that it represented a radical departure from another model of protein synthesis, whose importance and generality has generally not been recognized in the literature on the history of molecular biology (Olby 1970; Thieffry and Sarkar 1998; Morange 2000, chapter 12; Fruton 1999, chapter 8; Kay 2000, chapter 4; de Chadarevian 2002, chapter 6; Judson 1996, chapter 6).[1] By considering Crick's 1957 paper (Crick 1958) in this context, one can gain a better understanding of its historical significance and of its current meaning, as it still serves today as one of the intellectual foundations of the life sciences.

The Central Dogma

Francis Crick, a physicist-turned-biologist from Cambridge formulated the central dogma in a lecture on protein synthesis given at University College London in September 1957. He presented several hypotheses to account for a number of experimental facts that had recently been published. Most of Crick's claims were unoriginal. Following his tendency to interpret the work of others, he spelled out assumptions that colleagues had left implicit in their own work (Olby 1970, 976). He asserted, for example, that the specificity of a piece of nucleic acid resulted from the sequence of its bases (not its three-dimensional structure), and that 'this sequence is a (simple) code for the amino acid sequence.' He called this proposition the 'Sequence hypothesis', and believed it to 'be rather widely held' (Crick 1958, 152).[2] The second key idea he proposed was that 'once "information" has passed into proteins it *cannot get out again*' what Crick called the 'Central Dogma' (Crick 1958, 153). By 'information', Crick meant the precise order of

[1] However, H.F. Judson does emphasize the 'working out of the idea of specificity' as one-dimensional sequence (p. 582), but ignores how this idea replaced the alternative model discussed in this paper.

[2] Crick published a similar account a year earlier in *Scientific American*, where he discussed 'the hypothesis which my colleagues and I call the Central Dogma' (Crick 1957, 198).

the units along a nucleic acid or protein chain, i.e. its sequence. In other words, proteins did not influence the sequence of nucleic acids or other proteins; they contained information, but did not pass it on. Thus, nucleic acids were causally prior to protein, and in a certain way, more fundamental. This view reflected the then growing idea that DNA was the most important component of the cell, its 'master plan' (Gaebler 1956, 170; Keller 1995, chapter 2) or as geneticist George Beadle put it in 1957, 'a recipe for constructing a person' (Beadle 1957, 399).[3] What is perhaps most remarkable in Crick's views of protein synthesis is what it leaves out. For Crick, giving an answer to the problem of protein synthesis amounted to explaining how a protein acquired its amino acid sequence. On the subject of protein conformation, the three-dimensional folding up of the polypeptide chain, he remained almost completely silent. Given the fact that such diverse phenomena as the ability of hemoglobin to carry oxygen, of antibodies to recognize antigens, and of enzymes to carry out their catalytic activity all rested on the respective conformation of these proteins, how could Crick simply ignore this question? How could he claim to review the problem of protein synthesis without addressing the essential feature that made protein uniquely functional?

This intellectual step was made possible by making a bold assumption, namely that protein 'folding is simply a function of the order of the amino acids' (Crick 1958, 144). Thus, for a cell to make a specific protein, it was only necessary to specify its amino acid sequence, since its conformation and function would follow automatically. This key assumption simplified the problem of protein synthesis tremendously, as Crick recalled in 1970:

> Because it was abundantly clear by that time that a protein had a well defined three-dimensional structure, and that its activity depended crucially on this structure, it was necessary to put the folding-up process on one side, and postulate that, by and large, the polypeptide chain folded itself up. This temporarily reduced the problem from a three-dimensional one to a one-dimensional one. (Crick 1970, 561)

Crick could thus focus on the relationships between nucleic acid and protein sequences, speculating on the biochemical underpinnings (the role of RNA, the adaptor hypothesis) and the theoretical consequences (the coding question) of this problem. This assumption also gave the central dogma its empirical content. Indeed, protein sequences and conformations were tremendously difficult to

[3] Beadle acknowledged that it also required, in addition to an egg, 'some ten tons of food and a suitable environment' (p. 399).

determine experimentally, but protein activity could easily be monitored by a number of biochemical essays. Changes in sequences would then be inferred from changes in protein activity. Before examining how it became scientifically reasonable to make this key assumption, something Crick did not address in his recollection, I will briefly outline the reception of the central dogma, which will show how this assumption came to be embedded in the more popular understanding of the central dogma.

The Reception of the Central Dogma

The central dogma rapidly gained wide acceptance and by the mid 1960s, one is hard pressed to find any criticisms of its main ideas in the scientific literature.[4] Even though some of Crick's proposals were quite speculative when he presented them in 1957, they gained empirical support in the immediate following years. By 1961, a biochemist wrote in *Nature* that '[the Central dogma] is almost universally accepted' (Leslie 1961), and another one claimed that its core idea was 'so fundamental to present day thinking in the field of molecular biology that it has rightfully been referred to as the "central dogma"' (Mahler and Fraser 1961). A few years later, another researcher wrote in *The Lancet* that 'the Crick dogma is so authoritative that a good deal of experimental evidence would be needed to disestablish its all-embracing validity' (Field 1967). Another editorial published in 1967 noted that 'modern trinitarians have a deep faith in the Central dogma of molecular biology' (Anonymous 1967, 705).[5] Biochemist Felix Haurowitz, from Indiana University, was one of the rare dissenting voices. In the 1963 edition of his book entitled *Chemistry and Biology of Proteins*, he remarked that it would seem 'strange indeed that nature should not have made use of the transfer of information from protein to protein and also from protein to nucleic acids' (Haurowitz 1963, 438). The central dogma became more widely known when James Watson described it in his popular textbook, *Molecular Biology of the Gene*, published for the first time in 1965, a text that would train several generations of molecular biologists. For Watson, the central dogma amounted to a simple diagram, 'DNA' RNA'Proteins', or 'DNA makes RNA makes Proteins' (Watson 1965,

[4] The major exception being the biologist Barry Commoner, whose views will not be discussed here because they remained extremely marginal, see Commoner 1964, 3316-3317.

[5] They were called 'trinitarians' for their belief in the exclusive importance of DNA, RNA and proteins.

315),[6] a slogan that conveniently summarized the intellectual agenda of the new discipline.

In order to understand, beyond its sheer acceptance, the reception of the central dogma, one can examine episodes in which it was purportedly challenged, for moments of transgression reveal accepted norms, values and theoretical commitments. Reverse transcription, i.e. the production of DNA from an RNA template, and the replication of the scrapie agent, have represented two such challenges.

In 1970, Howard Temin and David Baltimore announced independently that they had isolated an enzyme, reverse transcriptase, that could copy RNA sequences into DNA sequences. An editorial in *Nature* claimed that this discovery had 'reversed the Central dogma' (Anonymous 1970), prompting Crick to publish a note restating that RNA-to-DNA information transfer was not excluded in his original scheme, but only in Watson's simplified version (Crick 1970; Darden 1995). The claims that information transfer from RNA to DNA challenged the central dogma were thus short lived.

The case of the agent causing the scrapie disease represented a much more serious challenge and the only one Crick worried about in 1970 (Crick 1970). In the 1960s, a number of experimental results pointed to the fact that the agent causing scrapie in sheep might be composed solely of proteins, not of nucleic acids, unlike all viruses and other pathogens (Keyes 1999a; Keyes 1999b). Thus, if this unusual agent replicated without the intervention of DNA or RNA, it would violate the central dogma, as several authors remarked, since information would be passed from proteins to proteins. However, this was not the only possible interpretation. The scrapie agent could also activate an existing gene in the cell, inducing its own synthesis, or change the conformation of preexisting proteins (Griffith 1967). Surprisingly, both of these cases have also been interpreted as violations of the central dogma. A comment published on the second interpretation in *The Lancet* in 1967 asserted that 'this would invalidate the accepted dogma of present day molecular biology in which D.N.A and R.N.A. control all biological activity' (Lewin 1972). An editorial about the third mechanism, published in *Nature Genetics* in 2002, claimed that 'Crick's original proposal [...] simply stated that information flow in the cell goes from nucleic acids to proteins. The obvious exception to this statement is the prion hypothesis, whose father, Stanley Prusiner, was awarded a Nobel Prize' (Anonymous 2002).

[6] In this scheme, Watson even omits the DNA to DNA transfer, which is however included on p. 298.

These comments reveal that the central dogma was understood in a much broader sense than either Crick or Watson had intended. The central dogma came to mean that nucleic acids were the sole determinant of protein specificity in the broadest sense of the word, i.e., not only protein sequence, but also protein conformation and activity, or as an editorialist put it bluntly, DNA controlled 'all biological activity' (Lewin 1972). Crick was careful to restrict his formulation to the definition of sequences and only assumed that these would determine the conformation and biological activity of proteins. This difference between Crick's statement and its reception points, once again, to the importance of this crucial assumption made by Crick, namely that protein conformation was determined by amino acid sequence. In order to understand how Crick could assume in 1957 that 'the polypeptide chain folded itself up', one needs to explore how protein formation was explained in the previous years. Linus Pauling was one of the most vocal advocates for a theory whereby proteins acquired their conformations from three-dimensional templates.

Pauling's Model of Biological Specificity

Linus Pauling was perhaps the leading American physical-chemist from the 1930s to the 1950s (Hager 1995). Chairman of the Division of Chemistry at the California Institute of Technology, he made essential contributions to explaining chemical properties, such as the nature of the chemical bond, in physical terms. In the late 1930s, he became increasingly interested in biological macromolecules, elaborating a molecular vision of life resting on the 'size and shape of molecules'.[7] This view was popular until the late 1950s and served as the theoretical framework to explain such diverse phenomena as the action of genes, the formation of antibodies, and the activity of enzymes.

Pauling began his incursion into biology by investigating the structural chemistry of the blood and the formation of antibodies. His approach to these different problems was guided, as he put it in a 1938 lecture, by his desire to account 'for the properties of substances in terms of the shapes of the molecules of which they are composed',[8] rather than by their ordinary chemical properties. In order to explain

[7] L. Pauling, typescript, 'Molecular Architecture and Biological Reactions', The George Westinghouse Centennial Forum, Pittsburgh, Pennsylvania, May 17, 1946, p. 7. Oregon State University, Special Collections, Eva and Linus Pauling papers (OSU Archives hereafter).

[8] L. Pauling, typescript, 'The Structural Chemistry of Blood', Pomona, California, March 10, 1938, p. 2, OSU Archives. See also later account such as (Pauling 1970).

the specificity of macromolecules – why enzymes only reacted with certain molecules for example – he developed the concept of 'complementarity', inspired by the German chemist Emil Fischer's 'lock-and-key' model of enzyme specificity formulated in the 1890s. For Pauling, specificity was essentially the result of the complementarity of two shapes, rather than their chemical composition. He gave a talk in 1946 that conveniently summarizes his approach:

> The specificity of the physiological activity of substances is determined by the size and shape of molecules, rather than primarily by their chemical properties, and [...] the size and shape find expression by determining the extent to which certain surface regions of two molecules (at least one of which is usually a protein) can be brought into juxtaposition - that is, the extent to which these regions of the two molecules are complementary in structure.[9]

The emphasis on complementarity was a consequence of Pauling's ideas about the nature of the physical forces at work in biology (Pauling and Delbrück 1940). For him, weak hydrogen interactions played an essential role in biological molecules (Pauling and Niemann 1939). Since the intensity of these forces diminish very rapidly with increasing distance, two molecules had to be complementary in order to present the largest possible surface to each other at the closest distance. It is hard to overestimate the importance of this idea in Pauling's thinking, since he applied it to virtually every problem in biology and medicine he addressed.

Instructive Theories of Antibody Formation

Pauling's research on antibodies brought empirical support to his views (Kay 1989; Morange 2000, chapter 12). Indeed, the complementarity principle was particularly effective in explaining antibody-antigen interactions, as the physician Paul Ehrlich, following Emil Fischer's lock-and-key model, had already suggested (Silverstein 2002). Pauling elaborated an 'instructive' theory of antibody formation, published in 1940, in which immunological specificity, i.e. the structure of the antibody, was acquired from the antigen which directed the folding of a peptide chain into a complementary structure (Figure 1; Pauling 1940).[10] For Pauling, 'all antibody molecules contain the same polypep-

[9] L. Pauling, typescript, 'Molecular Architecture', 1946, p. 7, OSU Archives.

[10] An instructive theory of antibody formation had already been proposed by Felix Haurowitz in 1930, but in his case, the antigen directed the sequence of the peptide chain (Breinl and Haurowitz 1930).

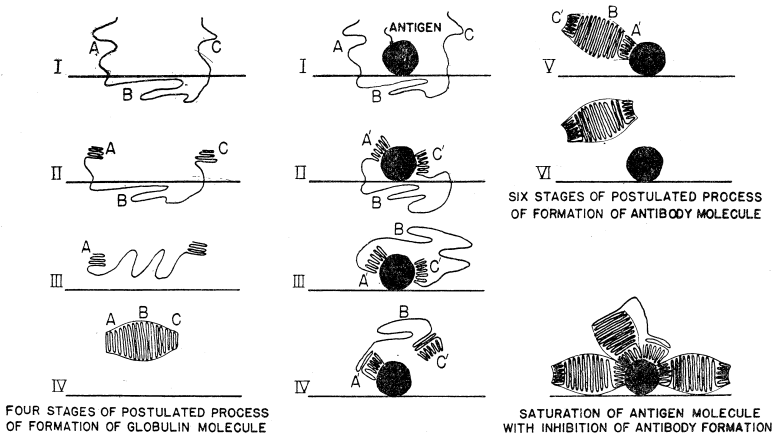

Fig. 1 - Diagrams representing four stages in the process of formation of a molecule of normal serum globulin (left side of figure) and six stages in the process of formation of an antibody molecules as the result of interaction of globulin polypeptide chain with an antigen molecule. There is also shown (lover right) an antigen molecule surrounded by attached antibody molecules or part of molecules and thus inihibited from further antibody formation.

tide chains as normal globulin, and differ from normal globulin only in the configuration of the chain; that is, in the way that the chain is coiled in the molecule' (Pauling 1940, 2644). These theoretical speculations soon gained additional experimental support when Pauling, together with the Caltech immunologist Dan H. Campbell, was able to obtain specific antibodies *in vitro* by mixing polypeptides and antigens (Pauling and Campbell 1942). Even though these *in vitro* experiments, first published in 1942, were not reproduced in other laboratories, Pauling held them to be correct until at least 1957.[11] Pauling's instructive theory of antibody formation was also supported by the results of several authors who found the sequence of different antibodies to be identical (Haurowitz 1956). Pauling's 'instructive' theory was the most widely accepted explanation of antibody formation, until its replacement by the clonal selection theory in the early 1960s (Burnet 1941; Burnet and Fenner 1949; Silverstein 1989, chapter 4).

Pauling did not think that this mechanism of protein formation was restricted to antibodies. He believed that 'the same process of molding of plastic materials into a configuration complementary to that of another molecule which serves as a template is responsible for all biological specificity'. In particular, he claimed that 'genes

[11] L. Pauling to F. Haurowitz, January 1, 1957, OSU Archives.

serve as the templates on which are molded the enzymes which are responsible for the chemical characters of the organism'.[12] In addition to gene action and antibody formation, Pauling believed that the principle of complementarity could explain gene replication, (Pauling and Delbrück 1940),[13] enzyme-substrate interaction, protein crystallization,[14] or even the occurrence of diseases such as hay fever.[15] In the following years, Pauling insisted in numerous speeches and papers on how much he hoped this principle would explain biological and pathological processes.

A number of other researchers used the example of antibodies to think about protein synthesis more generally. In 1945, Pauling's colleague at Caltech, the geneticist Sterling Emerson, had devised similar models to explain gene action and gene replication (Figure 2; Emerson 1945). Similarly, the biochemist Felix Haurowitz held

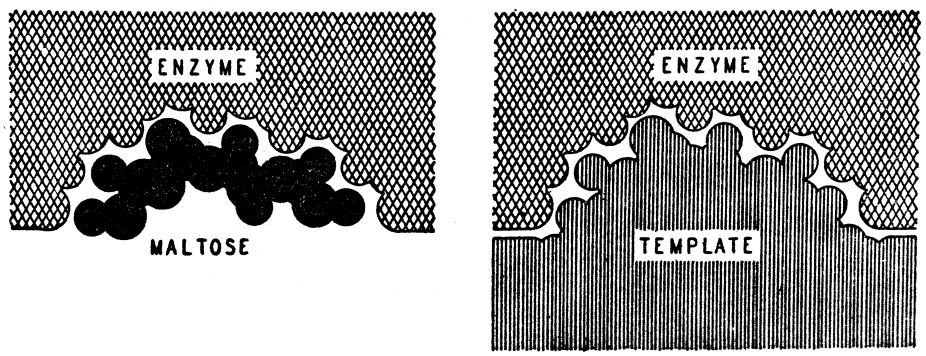

Fig. 2 - Complementary surfaces: left, section through maltase (ENZYME) with associated maltose molecule; right, maltose and complementary template.

almost identical views about the formation of antibodies and, like Pauling, used this case to generalize about protein synthesis. In the first edition of his book entitled *Chemistry and Biology of Proteins*, Haurowitz discussed the relation of amino acid sequence to protein structure:

[12] L. Pauling, *Molecular Architecture and the Process of Life,* Nottingham: Sir Jesse Boot Foundation, 1948, p 10, OSU Archives.

[13] L. Pauling, typescript, 'The Pick Lecture', Nu Kappa Nu Medical Fraternity, University of Chicago, September 3, 1945, OSU Archives.

[14] L. Pauling, typescript, 'The Interpretation of Some Chemical Properties of Hemoglobin in Terms of Its Molecular Structure', December 22, 1947, p. 12 and L. Pauling, typescript, 'The Structure of Antibodies and the Nature of Serological Reactions', 8 April 1948, OSU Archives.

[15] L. Pauling, typescript, 'The Significance of Structural Chemistry', George Fisher Baker Lectureship, Cornell University, October 12, 1937, OSU Archives.

> The synthesis of antibodies is a special case of protein synthesis, but it is certainly not different from the synthesis of the other proteins such as normal serum globulin. […] We have no reasons to assume two different types of protein synthesis. (Haurowitz 1950, 343-349)

The biochemist Richard L.M. Synge reviewed Haurowitz's book very favorably in *Nature*, in particular the view that 'three-dimensional folding' was determined by 'a variety of […] cellular constituent' (Synge 1952). It is important to emphasize that the question of protein synthesis was not restricted to the mechanisms by which a polypeptide acquired its conformation. Indeed, a polypeptide had to be synthesized in the first place. Thus, a number of researchers focused on the biochemical processes that led to the formation of the polypeptide chain. From the 1940s, a 'multi-enzyme' theory explained the synthesis of polypeptides by the action of many specific enzymes, often the same enzymes that were involved in protein degradation. The biochemist Joseph Fruton at Yale University, for example, was one of the main proponents of this view (Fruton 1941; Bartels 1983). Other researchers - for example, the cellular physiologists Jean Brachet and Torbjörn O. Caspersson in Europe (Brachet 1944; Caspersson 1950; Sapp 1987; Thieffry and Burian 1996)[16] and the biochemists Paul C. Zamecnik and Ernest E. Galein in the United States (Rheinberger 1997) - were more interested in the cellular location and the role of the different cellular constituents in protein synthesis, rather than the question of how proteins acquired their final conformation.

In the 1940s, empirical support for the three-dimensional template model of protein synthesis was scarce and mainly restricted to antibodies. The early 1950s brought new experimental evidence in favor of this model, arising from two very different experimental systems, enzymatic adaptation in bacteria and sickle cell anemia in man.

Experimental Support for the Three-Dimensional Template Theories

The phenomenon of 'enzymatic adaptation', later called 'enzymatic induction', was defined as the production of specific enzymes, usually in bacteria or yeast, following the addition to the media of an 'inducer',

[16] See also the special issue, edited by D. Thieffry and R. Burian of *History and Philosophy of the Life Sciences,* 19(1), 1997.

usually the enzyme's substrate. The French biochemist Jacques Monod, working at the Pasteur Institute in Paris, devised a number of experiments to understand how a small molecule could induce the production of enzymes. The results of different experiments led him to suggest, in 1945 and a number of times thereafter, that the inducer gave a preexisting polypeptide its three-dimensional structure and hence its specific enzymatic activity (Monod 1947; Monod and Cohn 1952; Gaudillière 1992; 1993). This mechanism was believed by Jacques Monod and others (such as the biochemist Henry Borsook and the microbiologist Martin R. Pollock), to be of general relevance for understanding protein synthesis (Borsook 1956; Pollock 1952).

Sickle cell anemia provided the most clear-cut example of how genes could specify the conformation of proteins, not their primary sequence. In 1945, when Linus Pauling learned about sickle cell anemia, he almost immediately thought that 'perhaps the Hb changes shape',[17] and set out to investigate whether sickle cell anemia was not due to an abnormal form of hemoglobin. In the summer of 1948, Harvey Itano, who had been hired by Pauling, was able to show that hemoglobin taken from sickle cell anemia patients differed in an electrophoresis apparatus from normal hemoglobin. The paper they published in *Science* a year later, provocatively entitled 'Sickle cell anemia, a molecular disease', made two important points (Pauling *et al.* 1949, Strasser 1999). First, it proposed a causal link between an abnormal molecule and its pathological consequences. Second, it showed for the first time that a gene could alter the physical properties of a protein and not just determine its presence or absence. In the following years, Pauling continued to investigate sickle cell anemia hemoglobin. This research is usually omitted in historical accounts, obscuring the role sickle cell anemia played in the debates about protein synthesis in the 1950s.

Pauling tried to generalize his sickle cell result to other 'molecular diseases', such as various blood diseases and mental diseases. But above all, he wanted to pinpoint the exact origin of the hemoglobin differences in sickle cell anemia. His grant requests to the Office of Naval Research in 1951, to the U.S. Public Health Service in 1954, and to the Ford Foundation in 1956, all stressed the importance of this question.[18] A difference in amino acid content would have easily explained the elec-

[17] L. Pauling, typescript, 'The Future of Medical Research', Talk For UMCA, California Institute of Technology, August 29, 1945, OSU Archives, 1945s.2; L. Pauling, typescript, 'The Pick Lecture', Nu Kappa Nu Medical Fraternity, University of Chicago, September 3, 1945, OSU Archives, 1945s.3.

[18] R. B. Corey to L. Pauling, January 1, 1951; Office of Naval Research 260.7; US Public Health Service Application, 1954; Ford Foundation, Mental Disorders 234.18, OSU Archives.

trophoresis results. As Francis Crick recalled, 'I was convinced (perhaps rashly) that there would be a change in amino acid composition.'[19] In 1950, the Caltech biochemist Walter A. Schroeder performed chromatographic analysis of amino acids, but did not find any difference which could explain the altered electrophoretic mobility (Pauling *et al.* 1950; Schroeder, Kay and Wells 1950). In a subsequent note published in *Science*, Pauling, Schroeder and other members of the Chemistry Division concluded that 'the electrophoretic differences of the hemoglobins could not be attributed to differences in [...] amino acid residues, but is presumably the result of [...] a difference in folding of the polypeptide chain'. (Pauling *et al.* 1950) This result was confirmed the same year by the same group and later by another (Schroeder *et al.* 1950; Huisman, Jonxis and van der Schaaf 1955). In order to examine by a more direct method whether sickle cell hemoglobin was folded differently, the immunologist Dan H. Campbell, Pauling's colleague at Caltech, set out to examine whether they reacted differently immunologically. As Pauling had expected, they did indeed, suggesting different conformations of the proteins (Goodman and Campbell 1953). These results gave additional support to the view that genes acted by directly defining the conformation of proteins.

As a result, in the mid-1950s, Pauling became more confident than ever that his views of protein synthesis were correct. In 1954, for example, he suggested in his Harvey lecture that normal and abnormal hemoglobin were 'composed of the same polypeptide chains, folded, however, in different ways'. He then outlined the possible genetic implications of this fact, namely that 'the gene responsible for the sickle cell abnormality is one that determines the nature of the folding of polypeptide chains, rather than their composition'.[20]

Francis Crick, on the other hand, was not so confident, as he recalled much later: 'Their method was in fact too crude to detect such a single change in amino acid composition. I clearly realized this at the time.'[21] Indeed, there were good reasons to question Pauling's and Campbell's negative results. Crick's colleague in Cambridge, the biochemist Frederick Sanger, had sought since the late 1940s to determine the exact sequence of a small protein, insulin (de Chadarevian 1996). Counter to what the fashionable Bergmann-Niemann theory of protein structure predicted (Olby 1979; Judson 1980; Morange 2000, chapter 12),

[19] Francis Crick to the author, July 20, 1998.

[20] L. Pauling, typescript, 'Abnormality of Hemoglobin Molecules in Hereditary Hemolytic Anemias', 1954, p. 8, OSU Archives.

[21] Francis Crick to the author, July 20, 1998.

Sanger's results showed convincingly that there was 'no simple periodic arrangement of residues along the chains' (Sanger 1952, 61). In the 1950s, it thus became clear that proteins had a specific sequence[22] which could represent a serious candidate for determining their very diverse conformations. However, Pauling, Haurowitz, Pollock, and others did not need to alter their views too much to incorporate Sanger's findings in their own schemes. They divided protein synthesis into a two-step process. In the first, genes played a crucial role by determining the sequence of polypeptides; in the second, gene products, inducers, or antigens, for example, folded the polypeptides in their final conformation (Borsook 1956; Haurowitz 1956). Similarly, the publication of James Watson and Francis Crick's double helix model of DNA in 1953, in which the sequence of nucleic bases determined the specificity of the nucleic acid, was readily incorporated in the three-dimensional template models of protein synthesis. DNA was postulated to be involved in specifying protein sequences, as Alexander Dounce had proposed a year before, or George Gamow just after Watson and Crick's paper (Gamow 1954; Dounce 1952).[23] Another template would then bring the polypeptide to assume its final configuration. The demise of the three-dimensional template model of protein synthesis did not come from the double helix, but from various other experimental breakthroughs.

The Demise of the Three-Dimensional Template Theories

In 1956, Pauling gave a series of speeches in Italy and France entitled 'Abnormal Hemoglobin Molecules in Relation to Diseases'.[24] The typescript of his speech contained the same views he had been advocating in the past years. However, probably sometime during his trip, Pauling added in handwriting in the margin, 'Recently (1956) Dr. Ingram in England has split hemoglobin enzymatically and found by chromatography that it forms 30 peptides; 29 are the same for HbA and HbS, and the 30th is different.'[25] Thus, it became very likely that sickle cell anemia hemoglobin had a different amino acid composition after all and thus there was no need to invoke differences in protein folding between the hemoglobins. The same year, Vernon Ingram, Crick's colleague in Cambridge,

[22] Judson gives a central role to this episode in the history of molecular biology, see in particular Judson 1993.

[23] For a discussion of Dounce's template model, see Campbell 1953.

[24] L. Pauling, typescript, 'Abnormal Hemoglobin Molecules in Relation to Disease', Rome, Italy, 1956, OSU Archives.

[25] L. Pauling, 'Abnormal Hemoglobin', 1956, typescript, unnumbered page. OSU Archives.

published his results, and a year later he was able to point to a single amino acid difference between normal and sickle cell hemoglobins explaining the difference in the electrophoretic pattern (Ingram 1956; Ingram 1957). Ingram concluded his paper by stressing that 'it is now possible to show, for the first time, the effect of a single gene mutation as a change in one amino acid of the hemoglobin polypeptide chain' (Ingram 1957, 326).

The theory of protein synthesis in which genes directed protein folding had just lost one of its most powerful arguments. Worse, Ingram's results supported the alternative theory where genes specified only amino acid sequences. The two other arguments in favor of the three-dimensional template model were also losing momentum at the exact same time. Indeed, the experimental results of David S. Hogness, Melvin Cohn, and Jacques Monod, showing in 1955 that enzymatic induction involved *de novo* synthesis of a protein, not the conversion of a precursor, cast some doubt on the idea that the inducer directed the specific folding of a polypeptide precursor (Hogness, Cohn and Monod 1955).[26] Similarly, the fact that enzymes could be induced by molecules that were not substrates, as Monod had outlined in a paper published in 1956 (Monod 1956), seemed to constitute convincing evidence that the inducer was not concerned 'with molding the enzyme configuration' (Schweet and Owen 1957, 203). At the 1957 conference where Francis Crick presented the central dogma, Martin R. Pollock, a long-time advocate of the idea that inducers served as templates in protein synthesis, concluded his presentation by 'it is clear that the cells already possess the "information" required to produce the specific protein' (Pollock and Mandelstam 1958, 200). When in the late 1950s, at the Pasteur Institute in Paris, the phenomenon of enzymatic induction started to be compared to lysogenic induction and not antibody formation, the inducer came to be understood as an element regulating the biosynthesis of proteins, not specifying protein folding.

The support from the experiments on antibody formation also began to fade in the late 1950s. In 1955, the microbiologist Niels K. Jerne proposed that the information for making antibodies was contained inside the cells producing them and did not come from the antigens (Jerne 1955; Söderqvist 2003, 175-190). This theory thus contradicted the current instructive theories of antibody formation. Two years later, David

[26] However, in that paper the authors did not completely abandon the idea that the relationship between the precursor and the active enzyme was similar to that between 'the normal and antibody globulin', p. 112. Monod was still considering that the inducer could play a role in enzyme folding three years later (Monod, 1958). For the decline of the inducer as template, see Lederberg 1956.

Talmage and Frank Macfarlane Burnet independently developed the clonal selection theory, in which the role of the antigen was solely to select antibodies of preexisting specificities, not to contribute to their determination by 'instructing' protein folding (Silverstein 1989). This theory explained several empirical findings, such as the presence of natural antibodies, that were not well accounted for in the instructive theory of antibody formation. Geneticist Joshua Lederberg rapidly drew the consequences of this new theory for the genetic control of protein synthesis, namely that each antibody had a unique amino acid sequence determined by a unique DNA sequence (Lederberg 1959).

Thus, if genes were only involved in determining the amino acid sequence of polypeptides, how did proteins fold into the exact conformation that gave them their unique specificity? The answer to this question came about through the work of biochemist Christian B. Anfinsen and his collaborators, working at the National Institutes of Health in Bethesda. In 1956, they denatured *in vitro* the enzyme ribonuclease, leading to the unfolding of the peptide chain, the breakage of its four disulfide bonds, and the loss of activity. When they left the polypeptide to refold itself spontaneously, not only did it recover its activity, but it reestablished the exact four disulfide bonds out of 105 possibilities, regaining its original conformation. Thus, there was no need to postulate anything directing protein folding beyond the mere amino acid sequence of the polypeptide chain. Their results were published in *Science* in April 1957 (Sela, White and Anfinsen 1957), five months before Crick's lecture. Their model of protein folding came to be known as the 'thermodynamic hypothesis', since it was only guided by thermodynamic laws, not by a three-dimensional template (Epstein, Goldberger and Anfinsen 1963).[27] Four years later, in 1961 at the Cold Spring Harbor Symposia on Quantitative Biology, Jacques Monod and François Jacob noticed one issue had not been discussed during the conference, 'evidently because it is implicitly considered as settled'. They were referring to the idea that '(non-genetic) *structural* information needed to be furnished [...] at the stage of tertiary folding in protein synthesis'. (Monod and Jacob 1961, 394)[28]

Thus when Francis Crick began to prepare his conference paper to be presented in September 1957, the picture of protein synthesis had just changed radically and Crick was well aware of the results outlined

[27] The authors concluded with confidence, 'there is no need to postulate the existence of a genetically determined template. The folding process is thermodynamically guided, and seems to lead inevitably, at least with small proteins, to the formation of a unique native configuration' (p. 447).

[28] Italics in original. The authors imply, wrongly, that this hypothesis was only considered for non-genetic information.

above, since they were produced by some of his closest colleagues and friends, such as Vernon Ingram and Jacques Monod, or published in the most visible scientific journals. The complex problem of how proteins acquired their three-dimensional structure could then be left aside for physical chemists to worry about. The molecular biologist, on the other hand, could concentrate on the mechanisms by which nucleic acids sequences determined protein sequences or, to use the jargon of the day, he could focus on the 'flow of information' (Crick 1958, 144).

Conclusions

Crick's formulation of the central dogma represented a turning point in the history of the life sciences. Even though most of its ideas were unoriginal,[29] it defined authoritatively a new intellectual agenda for the emerging discipline of molecular biology. With regard to protein synthesis and gene action, it restricted the problem to the understanding the 'flow of information', i.e. the transfer of one-dimensional sequence information between three types of molecules: DNA, RNA, and proteins. The historiography of molecular biology has described how this new template theory of protein synthesis replaced the multi-enzyme theory (Bartels 1983). However, this perspective obscures how Crick's model came to replace another theory, the three-dimensional template theory, which enjoyed wide currency until the late 1950s.[30] According to its main proponents, this view of protein synthesis remained speculative until its demise, but it lacked neither empirical support nor social authority, being endorsed by a number of eminent scientists and Nobel prize winners. By contrasting these two different template theories, this paper tried to highlight the significance of Crick's proposal, in particular the fact that it brought the problem of gene action and protein synthesis down to one dimension.

Up to the present day, much of the confusion about the central dogma came about because Crick used the term 'information' to replace 'specificity'.[31] He defined information narrowly as 'the precise determination of sequence'. By doing so, he conflated three notions of specificity (functional, structural, and sequential) that had remained distinct until then and replaced them by a single term, 'information'. As long as the three remained equivalent, using the

[29] See for example the contributions of Norman H. Horowitz and Joshua Lederberg in Gaebler 1956.

[30] On the relationship between these two templates theories see Morange 2000, 131.

notion of information did not raise any ambiguity. However, as soon as this equivalence broke down, as in the case of prions, the agent of scrapie disease, where proteins with a unique sequence assumed different structures and functions, the interpretation of the central dogma became confused. Any change in structural or functional specificity which was not related to protein and nucleic acid sequences came to be understood as an exception to the central dogma and the sequence hypothesis. Crick himself was very aware of the limitation of any explanatory scheme, including his own, focusing solely on nucleic acids sequences. As he put in a letter to Howard Temin in 1970, 'I do not subscribe to the view that all "information" is necessarily located in nucleic acids. The central dogma only applies to residue-by-residue [sequence] information'.[32] Thus, even for Crick, often heralded as a single-minded crusader for a DNA-centred vision of life, there was a broader intellectual agenda to pursue in molecular biology. Indeed, for Crick in 1970, 'the real question to ask is, how much extra information is required, in addition to DNA and the code, to make a particular cell work?' Crick did not believe he had answered that question in his famous lecture on protein synthesis given thirteen years earlier, as he reflected, unusually insecure, 'In the naive sense [the sequence hypothesis] clearly is true. In the highbrow sense it is very difficult to say exactly what it is.'[33]

Acknowledgments

I would like to thank Joseph Fruton, Brendan Matz, Michel Morange and Bill Summers for reading this manuscript, and the staff of the Special collections at Oregon State University for their precious help in making the Eva and Linus Pauling papers accessible. Partial funding for this research is from the Swiss National Science Foundation, project No. 105311–109973.

[31] As Joshua Lederberg would put it 'specificity' (or "information" as it is called nowadays)', Lederberg 1956, 167. Geneticist Guido Pontecorvo made the same association at the 1957 conference where Crick presented the central dogma, Pontecorvo 1958, 1.

[32] F. Crick to H. Temin, August 3, 1970, Profiles in Science, National Library of Medicine (profiles.nlm.nih.gov).

[33] F. Crick to H. Temin, Sept. 17, 1970, Profiles in Science, National Library of Medicine (profiles.nlm.nih.gov).

Abbreviations

OSU - Oregon State University, Special Collections, Eva and Linus Pauling papers.

References

Anonymous 1967, 'The Scrapie Agent', *Lancet*, 2: 705-706.
Anonymous 1970, 'Central Dogma Reversed', *Nature*, 226: 1198-1999.
Anonymous 2002, 'Wag the Dogma', *Nature Genetics*, 30: 343-344.
Bartels D., 1983, 'The Multi-enzyme Programme of Protein Synthesis - Its Neglect in the History of Biochemistry and its Current Role in Biotechnology', *History and Philosophy of the Life Science*, 5: 187-219.
Beadle G.W., 1957, 'The Genetic Basis of Biological Specificity', *Journal of Allergy*, 28: 392-400.
Borsook H., 1956, 'The Biosynthesis of Peptides and Proteins', *Journal of Cellular Physiology Supplement*, 47: 35-80.
Brachet J., 1944, *Embryologie chimique*, Paris: Masson & Cie.
Breinl F., Haurowitz F., 1930, 'Chemische Untersuchung des Präzipitates aus Hämoglobin und Anti-Hämoglobin-Serum und Bemerkungen über die Natur der Antikörper', *Hoppe-Seyler's Zeitschrift für Physiologische Chemie*, 192: 45-57.
Burnet F.M., 1941, *The Production of Antibodies*, Melbourne: Macmillan.
Burnet F.M., Fenner F., 1949, *The Production of Antibodies*, Melbourne: Macmillan.
Campbell P.N., Work T.S., 1953, 'Biosynthesis of Proteins', *Nature*, 171: 997-1001.
Caspersson T.O., 1950, *Cell Growth and Cell Function, A Cytochemical Study*, New York: Norton.
Commoner B., 1964a, 'Deoxyribonucleic Acid and the Molecular Basis of Self-Duplication', *Nature*, 203: 486-491.
Commoner B., 1964b, 'Roles of Deoxyribonucleic Acid in Inheritance', *Nature*, 202: 960-968.
Crick F.H.C., 1957, 'Nucleic Acids', *Scientific American*, 197: 188-200.
Crick F.H.C., 1958, 'On Protein Synthesis', *Symposia of the Society for Experimental Biology*, 12: 138-163.
Crick F.H.C., 1970, 'Central Dogma of Molecular Biology', *Nature*, 227: 561-563.
Darden L., 1995, 'Exemplars, Abstractions, and Anomalies. Representations and Theory Change in Mendelian and Molecular Genetics'. In: Wolters G., Lennox J.G. (eds), *Concepts, Theories, and Rationality in the Biological Science*, Pittsburg: University of Pittsburg Press, 137-158.
de Chadarevian S., 1996, 'Sequences, Conformation, Information: Biochemists and Molecular Biologists in the 1950s', *Journal of the History of Biology*, 29: 361-386.
de Chadarevian S., 2002, *Designs for Life. Molecular Biology after World War*

II, Cambridge: Cambridge University Press.
Dounce A.L., 1952, 'Duplicating Mechanism for Peptide Chain and Nucleic Acid Synthesis', *Enzymologia*, 15: 251-258.
Emerson S., 1945, 'Genetics as a Tool for Studying Gene Structure', *Annals of the Missouri Botanical Garden*, 32: 243-249.
Epstein C.J., Goldberger R.F., Anfinsen C.B., 1963, 'The Genetic Control of Tertiary Protein Structure: Studies With Model Systems', *Cold Spring Harbor Symposia on Quantitative Biology*, 28: 439-449.
Field E.J., 1967, 'The Scrapie Agent', *The Lancet*, 2: 1041-1042.
Fruton J.S., 1941, 'Proteolytic Enzymes as Specific Agents in the Formation and Breakdown of Proteins', *Cold Spring Harbor Symposia on Quantitative Biology*, 9: 211-217.
Fruton J.S., 1999, *Protein, Enzymes, Genes*, New Haven: Yale University Press.
Gaebler O.H. (ed), 1956, *Enzymes: Units of Biological Structure and Function*, New York: Academic Press.
Gamow G., 1954, 'Possible Relation between Desoxyribonucleic Acid and Protein Structure', *Nature*, 173: 318.
Gaudillière J.-P., 1992, 'J. Monod, S. Spiegelman et l'adaptation enzymatique. Programmes de recherche, cultures et traditions disciplinaires', *History and Philosophy of the Life Science*, 14: 23-71.
Gaudillière J.-P., 1993, 'Molecular Biology in the French Tradition? Redefining Local Traditions and Disciplinary Patterns', *Journal of the History of Biology*, 26: 473-498.
Goodman M., Campbell D.H., 1953, 'Differences in Antigenic Specificity of Human Normal Adult, Fetal, and Sickle Cell Anemia Hemoglobin', *Blood*, 8: 422-433.
Griffith J.S., 1967, 'Self-replication and Scrapie', *Nature*, 215: 1043-1044.
Hager T., 1995, *Force of Nature,* New York: Simon & Schuster.
Haurowitz F., 1950, *Chemistry and Biology of Proteins,* New York: Academic Press Inc.
Haurowitz F., 1956, 'The Nature of the Protein Molecule: Problems of Protein Structure', *Journal of Cellular Physiology Supplement*, 47: 1-16.
Haurowitz F., 1963, *Chemistry and Function of Proteins*, New York: Academic Press Inc.
Hogness D.S., Cohn M., Monod J., 1955, 'Studies on the Induced Synthesis of Beta-galactosidase in *Escherichia coli*: the Kinetics and Mechanism of Sulfur Incorporation', *Biochimica Biophysica Acta*, 16: 99-116.
Huisman T.H.J., Jonxis J.H.P., van der Schaaf P.C., 1955, 'Amino-acid Composition of Four Different Kinds of Human Haemoglobin', *Nature*, 175: 902-903.
Ingram V. M., 1956, 'A Specific Chemical Difference between the Globins of Normal Human and Sickle-cell Anemia Haemoglobin', *Nature*, 178: 792-794.
Ingram V. M., 1957, 'Gene Mutations in Human Haemoglobin: The Chemical Difference between Normal and Sickle Cell Haemoglobin', *Nature*, 180: 326-328.

Jerne N.K., 1955, 'The Natural-Selection Theory of Antibody Formation', *Proceeding of the National Academy Science*, 41: 849-857.
Judson H.F., 1980, 'Reflections on the Historiography of Molecular Biology', *Minerva*, 18: 369-421.
Judson H.F., 1993, 'Frederick Sanger, Erwin Chargaff, and the Metamorphosis of Specificity', *Gene*, 135: 19-23.
Judson H.F., 1996 [1979], *The Eighth Day of Creation*, New York: Cold Spring Harbor Press.
Kay L.E., 1989, 'Molecular Biology and Pauling's Immunochemistry: A Neglected Dimension', *History and Philosophy of the Life Science*, 11: 211-219.
Kay L.E., 2000, *Who Wrote the Book of Life. A History of the Genetic Code*, Stanford: Sanford University Press.
Keller E.F., 1995, *Refiguring Life, Metaphors of Twentieth-Century Biology*, New York: Columbia University Press.
Keyes E.M., 1999a, 'The Prion Challenge to the "Central Dogma" of Molecular Biology, 1965-1991. Part 1, Prelude to Prions', *Studies in History and Philosophy of Biological and Biomedical Sciences*, 30: 1-19.
Keyes E.M., 1999b, 'The Prion Challenge to the "Central Dogma" of Molecular Biology, 1965-1991. Part II, The Problem with Prions', *Studies in History and Philosophy of Biological and Biomedical Sciences*, 30: 181-218.
Lederberg J., 1956, 'Comments on Gene-Enzyme Relationship'. In: Gaebler O.H. (ed), *Enzymes: Units of Biological Structure and Function*, New York: Academic Press, 161-169.
Lederberg J., 1959, 'Genes and Antibodies', *Science*, 129: 1649-1653.
Leslie I., 1961, 'Biochemistry of Heredity: a General Hypothesis', *Nature*, 189: 260-268.
Lewin P., 1972, 'Scrapie: an Infective Peptide?', *Lancet*, 7753: 748.
Mahler H.R., Fraser D., 1961, 'Structural Mutations', *Nature*, 189: 948.
Monod J., 1947, 'The Phenomenon of Enzymatic Adaptation', *Growth*, 11: 223-289.
Monod J., 1956, 'Remarks on the Mechanism of Enzyme Induction'. In: Gaebler O.H. (ed), *Enzymes: Units of Biological Structure and Function*, New York: Academic Press, 7-28.
Monod J., 1958, 'An Outline of Enzyme Induction', *Recueil des Travaux Chimiques des Pays-Bas*, 77: 569-585.
Monod J., Cohn M., 1952, 'La biosynthèse induite des enzymes (adaptation enzymatique)', *Advances in Enzymology*, 13: 67-119.
Monod J., Jacob F., 1961, 'General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation', *Cold Spring Harbor Symposia on Quantitative Biology*, XXI: 389-401.
Morange M., 2000, *A History of Molecular Biology*, Cambridge: Harvard University Press.
Olby R., 1970, 'Francis Crick, DNA, and the Central Dogma', *Daedalus*, 99: 938-987.
Olby R., 1979, 'The Significance of the Macromolecules in the Historiography

of Molecular Biology', *History and Philosophy of the Life Science*, 1: 185-198.
Pauling L., 1940, 'A Theory of the Structure and Process of Formation of Antibodies', *Journal of the American Chemical Society*, 62: 2643-2657.
Pauling L., 1942, 'The Production of Antibodies *in vitro*', *Science*, 95: 440-441.
Pauling L., 1970, 'Fifty Years of Progress in Structural Chemistry and Molecular Biology', *Daedalus*, 99: 988-1014.
Pauling L., Campbell D.H., 1942, 'The Manufacture of Antibodies *in vitro*', *Journal of Experimental Medicine*, 76: 211-220.
Pauling L., Delbrück M., 1940, 'The Nature of the Intermolecular Forces Operative in Biological Processes', *Science*, 92: 77-79.
Pauling L., Itano H.A., Singer S.J., Wells I.C., 1949, 'Sickle Cell Anemia, a Molecular Disease', *Science,* 110: 543-548.
Pauling L., Niemann C., 1939, 'The Structure of Proteins', *Journal of the American Chemical Society*, 61: 1860-1867.
Pauling L. *et al.*, 1950, 'Sickle Cell Anemia Hemoglobin', *Science*, 111: 459.
Pollock M.R., 1952, 'Stages in Enzyme Adaptation'. In: Davies R., Gale E.F. (eds), *Adaptation in Micro-Organisms*, Cambridge: Cambridge University Press, 150-183.
Pollock M.R., Mandelstam J., 1958, 'Possible Mechanisms by which Information is Conveyed to the Cell in Enzyme Induction', *Symposia of the Society for Experimental Biology*, 12: 195-204.
Pontecorvo G., 1958, 'Self-reproduction and all that', *Symposia of the Society for Experimental Biology*, 12: 1-5.
Rheinberger H.J., 1997, *Toward a History of Epistemic Things, Synthesizing Proteins in the Test Tube*, Stanford: Stanford University Press.
Sanger F., 1952, 'The Arrangement of Amino Acids in Proteins', *Advances in Protein Chemistry*, 7: 1-67.
Sapp J., 1987, *Beyond the Gene: Cytoplasmic Inheritance and the Struggle for Authority in Genetics*, Oxford: Oxford University Press.
Schroeder W.A., Kay L.M., Wells I.C., 1950, 'Amino Acid Composition of Hemoglobin of Normal Negroes and Sickle-Cell Anemics', *Journal of Biological Chemistry*, 187: 221-240.
Schweet R.S., Owen R.D., 1957, 'Concepts of Protein Synthesis in Relation to Antibody Formation', *Journal of Cellular and Comparative Physiology*, 50: 199-228.
Sela M., White F.H., Jr., Anfinsen C.B., 1957, 'Reductive Cleavage of Disulfide Bridges in Ribonuclease', *Science*, 125: 691-692.
Silverstein A.M., 1989, *A History of Immunology*, New York: Academic Press.
Silverstein A.M., 2002, *Paul Ehrlich's Receptor Immunology: The Magnificent Obsession*, San Diego: Academic Press.
Söderqvist T., 2003, *Science as Autobiography: The Troubled Life of Niels Jerne*, New Haven: Yale University Press.
Strasser B.J., 1999, 'Sickle Cell Anemia, a Molecular Disease', *Science*, 286: 1488-1490.

Synge R.L.M., 1952, 'Facts and Speculations about Proteins', *Nature*, 170: 6.
Thieffry D., Burian R.M., 1996, 'Jean Brachet's Alternative Scheme for Protein Synthesis', *Trends in Biochemical Sciences*, 21: 114-117.
Thieffry D., Sarkar S., 1998, 'Forty Years under the Central Dogma', *Trends in Biochemical Sciences*, 23: 312-316.
Watson J.D., 1965, *Molecular Biology of the Gene*, New York: Benjamin.