



HAL
open science

Proteins as monads

Sonia Dheur, Sven J Saupe

► **To cite this version:**

Sonia Dheur, Sven J Saupe. Proteins as monads. Peter Fratzl, Michael Friedman, Karin Krauthausen and Wolfgang Schäffner. *Active Materials*, De Gruyter, pp.339-366, 2022, 978-3-11-056181-4. <10.1515/9783110562064-016>. <halshs-02283978>

HAL Id: halshs-02283978

<https://shs.hal.science/halshs-02283978v1>

Submitted on 4 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

Sonia Dheur, Sven J. Saupe
Proteins as Monads

“As the wind of time blows into the sails of space, the unfolding of the universe nurtures the evolution of matter under the pressure of information.”¹

“Placed on the opaque canvas, these folds, cords, or springs represent an innate form of knowl - edge, but when solicited by matter they move into action.”²

1 Introduction

Active matter deals with elementary chemical and biological components and their collective properties-among others, macromolecular cell constituents capable of converting chemical energy into motion [Sanchez et al. 2012]. It is proposed that “the field of active matter is concerned with non-equilibrium systems in which the individual, constitutive units are themselves internally driven: machines made from machines” [Needleman and Dogic 2017, p. 1]. We would like to recall here that the formation of these elementary biological machines is – in itself – an internally driven process. If active matter indeed defines materials “able to change their shape in response to changes in environmental conditions and thus perform mechanical work on the nano-, micro- and macroscales” [Ionov 2014, p. 5015], then proteins in general, viewed as the Leibnizian machines that make up other organic machines [Leibniz 1991],³ are active materials. Proteins, regardless of their biological function, have

1 [Lehn 2002, p. 4763].

2 [Deleuze 2006, p. 4]. “Ces plis, cordes ou ressorts constitués sur la toile opaque, représentent les connaissances innées mais qui passent à l’acte sous les sollicitations de la matière” [Deleuze 1988, p. 6].

3 Leibniz defines organic bodies as machines made of machines. “Thus the organic body of each living being is a kind of divine machine or natural automaton, which infinitely surpasses all artificial automata. For a machine made by the skill of man is not a machine in each of its parts. [...] But the machines of nature, namely, living bodies, are still machines in their smallest parts *ad infinitum*” [Leibniz 1898, p. 254]. (“Ainsi chaque corps organique d’un vivant est une espèce de machine divine, ou d’un automate naturel, qui surpasse infiniment tous les automates artificiels. Parce qu’une machine faite par l’art de l’homme, n’est pas machine dans chacune de ses parties [...]. Mais les machines de la nature, c’est-à-dire les corps vivants sont encore des machines dans leurs moindres parties, jusqu’à l’infini” [Leibniz 1991, pp. 161–162].) After Descartes, the use of the word “machine” to designate an organic body is common, but the word is used loosely with very different meanings and intentions and sometimes even without an explicit underlying philosophical thesis [Andraut 2011]. In the dualist view of Descartes, the organic body is presented as a machine in order to define it as separated and subordinated to the soul, While the Leibnizian perspective of machine made of machines implies a radically distinct philosophical meaning and incorporates the

the inherent property to set themselves in motion, to fold into their native structure. The notion of motion often embedded explicitly in the active matter concept does not take here the form of a self-propelled displacement of an object in three-dimensional space but rather corresponds to a change of shape, a folding, that nonetheless corresponds to a movement. This biological self-organization principle, here at the most basic level, is then followed by various levels of supramolecular self-organization of increasing complexity, leading to the Leibnizian view of a living machine made of machines. We will insist here on the relation of genetic information to protein structure, to describe proteins as instructed components or informed matter that attains a determined shape (and supramolecular-organization state) thanks to the information contained in their primary sequence and in response to their environment [Dill and MacCallum 2012]. Protein folding envisioned as the most basic form of biological development and morphogenesis simultaneously contains elements of preformism and epigenesis.⁴ After a discussion of general aspects of protein folding and self-assembly and the interconnected role of the genetic and folding/assembly code in these processes, we will describe how proteins can display emergent properties through their collective behavior in the form of phase transition phenomena, and turn to the specific example of prion proteins to depict how in a further step of supramolecular self-organization biological matter can store and convey genetic information. Protein activity is classically framed into a structure-function paradigm which gradually gets replaced by a plasticity-function concept that takes into account how proteins react to their environment by changing shape. To think

notion of a machine animated from within. The materialistic monism of Julien Offray de La Mettrie, with *l'homme machine*, rejects that dualistic view of Descartes and the notion of a separate soul that would escape mechanism. His intention is to materialize the soul rather than to spiritualize matter. The use of the term “desiring machines” by Gilles Deleuze and Félix Guattari is closer to the Leibnizian sense as it confers a level of autonomy to the machines, the very wording (desiring machines) implies that machines have an internal agenda. For Deleuze and Guattari, the word also applies to social bodies and is not used in a theological context, in contrast to the Leibnizian perspective in which organic bodies are presented as divine machines.

⁴ These opposing concepts traditionally structure the debate on animal embryology and biological development in general [Van Speybroeck, De Waele, and Van De Vijver 2002]. Preformism envisions development as the growth and unfolding of preexisting structures already present in the egg, spore, or sperm and is traditionally associated to a theological dimension as the preexisting structures are presented as implanted by divine intervention. Epigenesis views development as the differentiation of new structures from homogenous matter lacking preexisting organization and philosophically emphasizes an autonomous nature of organic matter. Deleuze in *Le pli (The Fold)* finds that preformism and epigenesis both conceive the organism as a fold (“An organism is defined by endogenous folds” [Deleuze 2006, p. 7] (“Un organisme se définit par des plis endogènes” [Deleuze 1988, p. 11])) and makes a direct reference to protein folding. The modern conception of protein folding is neither purely preformist as protein structures do not grow out of preexisting analogous structures nor purely epigenetic as the protein contains primary structure (the sequence of amino acids) and in that sense is not homogenous and undifferentiated.

of these basal biological processes, we follow the Deleuzian reading of Gottfried Wilhelm Leibniz [Deleuze 1988] and propose that protein could be envisioned as *monads* and recall that Gilles Deleuze and Félix Guattari have explicitly defined proteins as *machines désirantes (desiring machines)* [Deleuze and Guattari 1972]. Throughout, we will call upon idiosyncratic examples because biological phenomena, in spite of all categorization efforts, often tend to bend and break constructed theoretical frames [Bergson 2013 (1907)] and are maybe best approached in the style of the schizo-analysis promoted by Deleuze and Guattari [Deleuze and Guattari 1980]. We will therefore attempt to combine philosophical and biological perspectives to review certain aspects of protein biochemistry.

2 The Protein Folding Problem, Anfinsen's Principle

Most of the structural and chemical work in living organisms is carried out by proteins. Proteins are central biological agents. The rules of the complex stereochemical templating procedures establishing the relation of the DNA sequence of genes to the amino acid sequence of the corresponding proteins have been laid out in the 1970s and termed the genetic code.⁵ But there is a second code, a folding code. After their mere material synthesis, to develop into functional entities, proteins have to fold into their *native* conformation. It is through this process of folding that the genetic information acquires biological meaning. The so-called Anfinsen principle states that the information specifying the fold is contained in the protein sequence, hence in the DNA sequence of the corresponding gene.⁶ The experimental basis behind the Anfinsen principle is deceptively simple and rests on the fact that protein denaturation (chemically induced unfolding) is reversible *in vitro*. Take an active purified protein, for instance a nuclease (an enzyme degrading nucleic acids) from *Staphylococcus* (Fig. 1A). In the presence of

⁵ The book by Lily Kay has fully described the historical conditions of emergence of the code metaphor for the protein translation process [Kay 2000].

⁶ It is of interest here to cite directly from Anfinsen's Nobel lecture of 1972: "the three-dimensional structure of a native protein in its normal physiological milieu (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and other) is the one in which the Gibbs free energy of the whole system is lowest; that is, that the native conformation is determined by the totality of interatomic interactions and hence by the amino acid sequence, in a given environment. In terms of natural selection through the 'design' of macromolecules during evolution, this idea emphasized the fact that a protein molecule only makes stable, structural sense when it exists under conditions similar to those for which it was selected – the so-called physiological state. [...] we can predict, in advance, the three-dimensional, phenotypic consequences of a genetic message" [Anfinsen 1973, pp. 223, 229]. These citations emphasize the articulation of information to meaning that is embedded in the protein folding process and stress that the conversion of information to meaning does not occur at the step of the decoding of the triplet information into amino acid by way of the genetic code but during the active folding of the protein.

high concentrations of the chemical urea, the protein is *denatured* (i.e., loses activity and fold), but when the chemical is removed, it spontaneously re-acquires activity and native fold, the protein *renatures*. It follows that the molecule contains endogenously the information to set itself in motion to fold into its native structure.⁷

Following Anfinsen's principle, of the astronomical number of shapes a nascent protein chain can adopt, the native fold is the one of lowest energy. Energetically, protein folding is a downhill process occurring spontaneously with a change in free energy in the range of 5–10 kcal/mol. The formation of novel non-covalent chemical interactions in the native fold has a positive contribution to folding while the decrease in entropy of the protein chain negatively contributes to folding. One of the main drivers of folding is the so-called hydrophobic effect leading to a collapse of the protein chain that secures hydrophobic amino acid side chains from the energetically unfavorable interaction with water molecules. In the general case, folding roughly corresponds to the transition from an extended highly mobile state to a more compact globule-like shape, with extreme diversity and variation from protein to protein.⁸ The protein folding process is often represented graphically as a protein transiting downhill in a funnel or landscape to a lower free energy state and from many conformations to a unique native fold (Fig. 2).

Ultimately, most forms of biological motion, as performed by biological nanomotors, are underpinned by conformational changes in proteins; in these cases, however, force generation is coupled with the expenditure of chemical energy (often ATP) [Guo et al. 2016]. To convey the general impression that proteins set themselves into

⁷ The semantics of protein folding embeds the dualistic preformist and epigenetic aspects of protein morphogenesis. The protein is synthesized as a *nascent* chain which then folds into its *native* state (in *native* conditions), can be *denatured* by heat or chemicals, and can further *renature* if conditions are native anew. This formulation implies paradoxically that although the fold is designated native (inborn) it is not present in the newborn (nascent) protein and its native (inborn) character is determined by interaction with the environment (the native or denaturing conditions). The unfolding process deprives the protein of its nature, but this lost part may be regained, the protein is equipped with the information to re-nature itself from within, given the right conditions.

⁸ In a paper founding the field of structural biology of proteins and describing for the first time X-ray crystallography 3D protein structure, the authors note that “perhaps the most remarkable features of the molecule are its complexity and its lack of symmetry. The arrangement seems to be almost totally lacking in the kind of regularities which one instinctively anticipates, and it is more complicated than has been predicted by any theory of protein structure. Though the detailed principles of construction do not yet emerge, we may hope that they will do so at later stages of analysis” [Kendrew et al. 1958, p. 665] (Fig. 1B). These words are in sharp contrast with those spawned by the discovery of the structure of DNA. For instance, Max Delbrück stated “The whole business was like a child’s toy that you could buy at the dime store, all built in this wonderful way that you could explain in *Life* magazine so that really a five-year-old can understand what’s going on [...] This was the greatest surprise for everyone” [Judson and Delbrück 1996, p. 41]. While DNA is immediately intelligible, proteins are not. DNA is most often referred to in the singular – against rationality – and proteins in the plural.

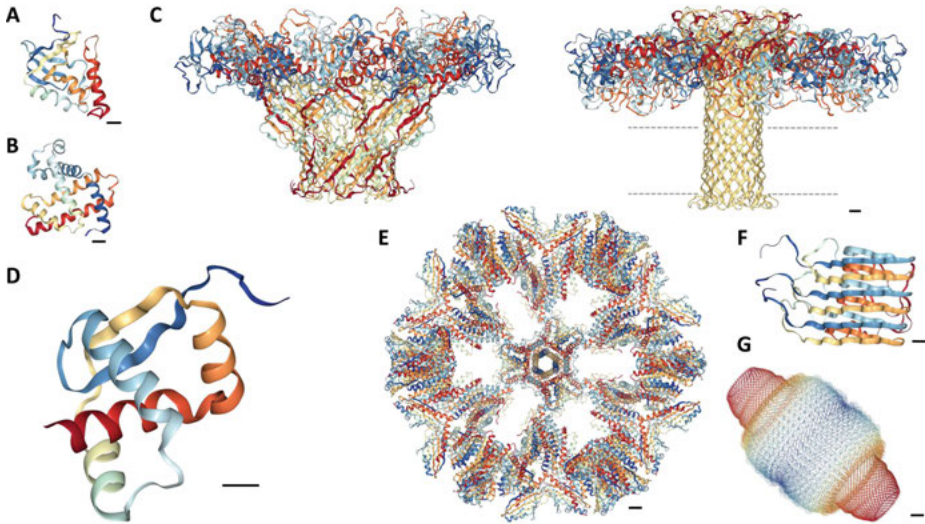


Fig. 1: Protein folds. The figure depicts the folds of different natural and designed proteins. For each protein, the name of the PDB file (protein database) containing the coordinates of all atoms of the protein and which was used to generate the protein image is given in parentheses: (A) Nuclease from *Staphylococcus aureus* (pdb, 2 sns) the protein used in the classical experiments of Christian Anfinsen on the general principles of protein folding. (B) Myoglobin, the first protein structure that was solved by X-ray crystallography, here the myoglobin of the sperm whale (*Physeter macrocephalus*) (pdb, 1mbn). (C) Two conformations of aerolysin, a pore-forming toxin produced by the human pathogen *Aeromonas hydrophila* capable of opening pores in cell membranes and causing cell death. The left image corresponds to the pre-pore state (pdb 5jzh), the right image to the pore state (pdb 5jzl) producing membrane damage. The dotted lines depict the position of the two membrane leaflets. (D) Protein MJ0366 of unknown function of the archaea *Methanocaldococcus jannaschii* (pdb 2 efv) as an example of a protein forming a knot, in this case a trefoil knot. (E) A designed protein complex forming a icosahedral cage, a polyhedron with 20 faces that has been produced by the group of David Baker (pdb 5im6). (F) The HET-s prion of the filamentous fungus *Podospora anserina* (pdb 2kj3), the first solved prion structure. (G) Major vault protein from *Rattus norvegicus* (pdb 4V60), this protein assembles into 78 subunit cage-like structure found inside cells and whose exact function remains unclear. Scale bar is 0.5 nm in (A)–(D) and (F), 1 nm in (E) and 5 nm in (G). All images were generated by the authors (S.D./S.S.) with Mol* at <http://www.rcsb.org> (Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, Velankar S, Burley SK, Koča J, Rose AS. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.* Accessed May 6, 2021).

motion during folding (even without ATP expenditure) a few idiosyncratic examples can be of use. A class of bacterial proteins termed autotransporters uses the energy of folding of a specific protein domain (misnamed passenger domain) to transport virulence factors across the membrane. This transport is ATP-independent and energy for the transport is provided by the folding of the passenger domain [Renn et al. 2012]. So, in this specific example, folding is translated to displacement of the protein in space

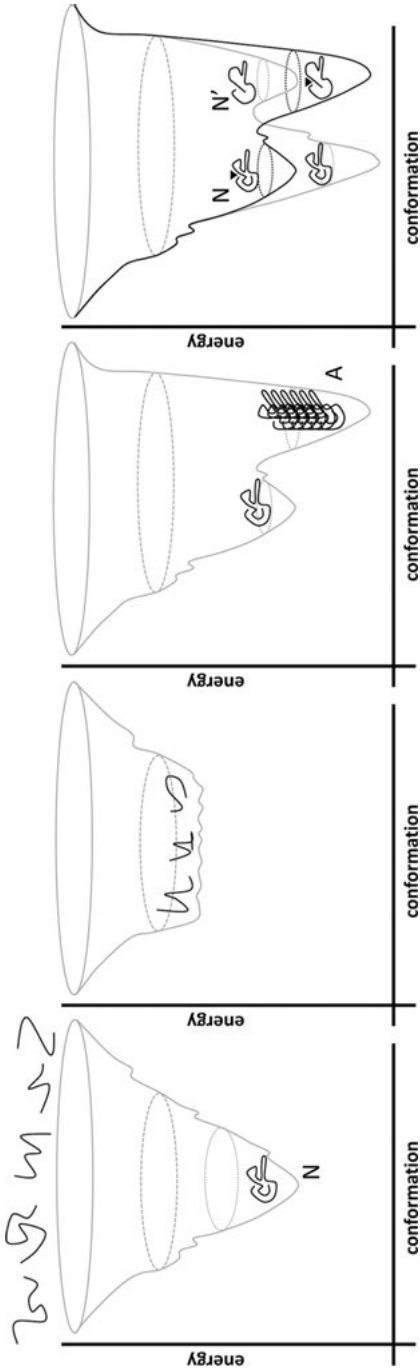


Fig. 2: Protein folding landscapes. The first panel on the left corresponds to the classical representation of folding as a protein folding funnel or protein folding landscape where the horizontal axis represents conformations and the vertical axis energy. The protein, represented as a solid curved line, exists as many different conformation displaying similar energies (at the top of the funnel). During the folding process, the protein acquires a unique conformation and a state of minimal free energy (at the bottom of the funnel), the native state marked N. This diagram is an illustration of protein folding process as described by Christian Anfinsen. The second panel represents the protein folding landscape in the case of IDPs. In that case, there is no single energy minimum corresponding to a native folded state, instead the protein exists in a plethora of interconverting conformations with comparable energy level, leading to this representation of a funnel with a flattened bottom. The third panel depicts the case of proteins capable of supramolecular assembly such as for instance proteins forming amyloid polymers. In this case, in addition to the energy minimum corresponding to the native state (as in the first panel) there is a second energy well of even lower energy corresponding to proteins assembled into higher order polymers (marked A). The assembled protein may or may not maintain the native conformation in the polymeric form. The last panel illustrates the process of allostery by which a protein can exist in two different conformations depicted by the different geometry of the solid curved line and marked N and N' and the relative free energy of the two conformations varies depending on the physico-chemical environment, more specifically by the presence or absence of a ligand (represented by the black triangle). In the absence of the ligand (gray funnel), the N conformation is the one of lower energy and is adopted, while in the presence of the ligand (black funnel) the N' conformation is favored. This process illustrates how sensing of the physico-chemical environment translates to a conformation change in the protein. Sensing of the ligand causes the protein to change from the N to the N' conformation. © Graphic by the authors (S.D./S.S.).

and through a membrane compartment. One can also mention the case of bacterial pore-forming toxins such as aerolysin, in which the re-folding of a region of the protein is able to poke holes into the membrane of the host cell under attack [Bokori-Brown et al. 2016] (Fig. 1C).⁹ Finally, to illustrate the complexity of the movements involved in the folding process, one might recall that about 1% of the proteins of known structure are not only folded but knotted; the functional roles of such knots remain poorly understood and cover a range of different topologies from twist knots that can be tied or untied with a single threading to more complex slipknots [Jackson, Sunna, and Micheletti 2017] (Fig. 1D).

3 So Many Shapes, the Levinthal Paradox

Anfinsen's principle comes with a paradox. The nascent or unfolded protein exists under astronomical numbers of interconverting conformers and gradually acquires a unique defined native fold with a decrease in entropy. Folding is thus a path from an extreme disordered to a unique ordered state. Each protein molecule experiences a transition from chaos to order, its own private cosmogony. Cy Levinthal pointed to the fact that the number of degrees of freedom in polypeptide chains is such that it would take longer than the age of the universe for a protein to sample all possible conformations until the native one is reached. Yet this process takes only microseconds to seconds depending on the protein. This so-called Levinthal paradox can be resolved by supposing that folding occurs through specific pathways that gradually lead the protein into the native fold, *Holzwege* are thus avoided. Currently, two opposing views on how exactly the Levinthal paradox is solved exist. The first hypothesis proposes that discrete subportions of the proteins of about 25 residues termed 'foldons' gradually take shape in a hierarchical order and that native folding occurs through assembly of such 'foldons' through a discrete, defined pathway (in a divide-and-conquer or local-to-global fashion) [Englander and Mayne 2017]. The opposing view supports the existence of a multitude of folding pathways and is described as the energy landscape theory in which the folding pathways are represented graphically as multiple possible displacements of the protein chain along a funnel-like landscape in which the *x*- and *y*-axes represent conformational space and the *z*-axis the decrease in free energy [Eaton and Wolynes 2017] (Fig. 2). In the first model, folding occurs in a defined pathway, in the second there are multiple possible folding pathways leading

⁹ Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein (<https://www.youtube.com/watch?v=ZOBtKZcePZo>, accessed November 26, 2019).

to the native state. This is no minor disagreement [Baldwin 2017], so that, in a way, the protein folding problem still stands.¹⁰

The Levinthal paradox in fact extends to the next step of molecular self-organization, at the level of supramolecular assembly into higher order complexes. Some emphasize that by analogy to protein folding the formation of the cellular interactome, that is, the higher order assembly of all the cellular macromolecular complexes (machines made of machines), cannot be achieved by a simple combinatorial trial-and-error process but occurs following ordered pathways that are collectively encoded in genomes and guided by preexisting cellular structures. It follows that a cell can only be assembled from a preexisting cell [Tompa and Rose 2011]. This concept of preexisting structural and super-structural information that every cellular component is supplied with evokes a form of preformism distinct from the strictly gene-centered view; giving a new meaning to the *omnis cellula e cellula* statement enunciated by Rudolf Virchow in the formulation of the cell theory [Virchow 1855, p. 311]. This extension of the Levinthal paradox at the level of the interactome also represents a critique of the performances claimed by the field of synthetic biology. The so-called synthetic bacteria engineered by the group of Craig Venter do indeed contain a synthetic genome but were obtained by introducing this genome in a preexisting natural cell.

4 Chaperones, Protein Memory, and Silent Mutations, Some Refinements of Anfinsen’s Principle

Like many ‘laws’ in biology, Anfinsen’s principle suffers exceptions, yet it serves as a general background rule onto which particular cases can be contrasted. If folding is fundamentally viewed as an intrinsic, immanent process, in practice, many proteins are difficult (to say the least) to ‘renature,’ that is, to fold *de novo* in vitro. In the biological context, folding often relies on the assistance of so-called chaperone proteins that ensure that folding proceeds normally in the highly crowded cellular environment. Such chaperones can also be required for the supramolecular assembly steps as mentioned in the previous section. The chaperone metaphor was put forward to stress that such proteins function to “prevent incorrect [...] interactions” [Laskey et al. 1978, p. 420]. One typical example of such chaperone proteins is the bacterial GroEL/GroES complex which functions as an isolation chamber fitting a

¹⁰ Molecular dynamics simulation now allows modeling the folding pathway at least for small fast folding domains that fold in the nano- to microsecond range. See, for instance, simulation of millisecond protein folding: NTL9 (from Folding@home) (<https://www.youtube.com/watch?v=gFcp2Xpd29I>, accessed November 26, 2019).

single nascent protein chain that then folds shielded from the multiple interactions encountered in the cytoplasm and exits the chamber once folded, as if execution of the intrinsic development program of the protein chain requires isolation to be completed.

Another interesting quasi-violation of Anfinsen's principle is provided by the phenomenon of protein memory [Shinde, Liu, and Inouye 1997]. Certain proteases (proteins that degrade proteins) are expressed as precursors with a so-called propeptide. This propeptide binds to the enzyme and inhibits its activity but is then cleaved off by self-maturation (the protein cleaves itself to eliminate the propeptide and attains full activity). Without this propeptide, however, the protein cannot fold and mutations introduced in the propeptide region affect the folding state of the protease, so that two proteases with the exact same sequence but having interacted with distinct propeptides acquire slightly different folds and have different activities. This process has been termed protein memory because the protein fold is shaped by past interactions and not only by the present conditions.

It is also becoming increasingly apparent that proteins are dependent on information contained in the triplet codons that is distinct from the specification of the amino acid type. The genetic code is termed degenerated because different codons (so-called synonymous codons) can encode the same amino acid, and mutations that do not alter the amino acid sequences are therefore termed silent (and often correspond to those occurring on the third position of the codons). It turns out that in many cases silent mutations are not, and alter the meaning of the genetic message. This is because different codons (encoding the same amino acid) lead to different rates of translation and the rate of translation or pauses in the translation process can alter the folding pathway of nascent proteins. So fold and function can be modified, even though the sequence of the protein is unchanged. For instance, *Neurospora* mutants bearing only silent (synonymous) mutations in the *frq* biological clock gene lose circadian rhythmicity [Zhou et al. 2013]. So, not unlike synonyms in natural language, no two codons are actually perfectly synonymous. Through this additional layer of informational content the genetic code and the folding code are intimately interwoven. In contrast to the genetic code, the folding code, this code within the code, is not yet fully cracked.¹¹

¹¹ This latter code-script derived metaphor is not well suited for a heuristic description of the folding code, that is probably best understood as governed by a rule-code. It has been stated that the term “code-script” appearing in Erwin Schrödinger's *What is life?* of 1944 was a sort of prescient forerunner of the discovery of the genetic code. But Walsby and Hodge argue that Schrödinger's code is to be understood as a rule-code rather than a cipher-code and could therefore apply to other aspects of biological development such as protein folding rather than being restricted to the cipher-code-like character of the genetic code [Walsby and Hodge 2017].

5 Protein Design, from *Machines Désirantes* to *Machines Désirées*

In protein science, there is currently a marked shift of emphasis from causal explanation to empirical prediction, a trend common to different fields of knowledge.¹² The main aim is no longer to explain how proteins fold but to predict fold from sequence and, most of all, to devise sequences with desired folds and functions. Active design now overshadows passive (naturalistic) description of the folding process. Structure prediction methods have gradually acquired better accuracy thanks to refinements in the energy functions used to calculate the minimal energy states, progress in available computational power and increase in the number of reference experimentally obtained protein structures. Protein structure prediction is also fueled by the gigantic amount of protein sequence produced by genomic and metagenomic approaches and it is now possible to identify many novel protein folds in metagenomics data sets corresponding to microbial organisms that are otherwise unknown and unculturable [Ovchinnikov et al. 2017]. A critical improvement was also the use of contact prediction based on residue co-evolution. Evolutionary information is implemented in the prediction scheme by analyzing residue co-variation in a large number of homologs of the target protein from different species. Residue pairs which co-vary during evolution are predicted to be close in space and to form contacts in the folded protein. The collective information contained within the protein sequences of homologs (which are the molecular descendants of an ancestral protein) and resulting from evolutionary trajectories spanning often millions of years in hundreds of different species provide the structural map of the protein shape. Shape is enciphered in history and history reveals the present shape. Total ab initio structure prediction (without help of any type of evolutionary information) is still out of reach in most cases.¹³

12 Michael Scriven noted that however ‘enlightening,’ causal explanation is doomed to always remain partial and a never-ending endeavor and, following Jean-François Lyotard, it may be possible to trace back the deflation of the valence attributed to explanation to the wreck of the great meta-narratives of modernity [Scriven 1959; Lyotard 1979].

13 From the point of view of the sociology of science, the protein fold prediction issue spawned several original initiatives. The CASP competition (for Critical Assessment of protein Structure Prediction) is a community-wide protein structure prediction experiment held every two years in which the different competing labs critically assess the performance of their prediction methods and software in a double-blind setting using a list of proteins whose actual experimentally determined structures are unknown to the organizers and the predictors [Kryshtafovych et al. 2018]. The protein folding field also led to different projects of citizen-science, like rosetta@home, in which more than a million participants contribute spare computing time on their personal computer or smartphone to the calculation intensive functioning of the Rosetta protein folding energy prediction algorithm. Likewise, the Foldit software (implemented also as foldit@home), is a citizen-science video game in which the participants solve heuristically protein folding problems [Cooper et al. 2010]. The latest

On the side of protein design, David Baker, head of the Institute of Protein Design at the University of Washington, is certainly the uncontested leader. His group could successfully design *de novo* proteins with a variety of folds, types of supramolecular organization and desired functions, including, in the most sophisticated case, a synthetic virus-like protein capsule capable of incorporating its own RNA genome [Butterfield et al. 2017] (Fig. 1E). The sequence space that can be explored in *de novo* protein design is proposed to be much larger than the actual sequence space explored by proteins from extant organisms (which is compared to an “archipelagoes in a vast sea of unexplored territory” [Huang, Boyken, and Baker 2016, p. 321]). In an unrestrained style, the advent of *de novo* protein design is described by this group as analogous to the transition from the Stone to Iron Age. A promotion video explains, “Early humans adapted objects they found in their surrounding environment to solve the challenge they faced, modern man, on the other hand, has learned to solve problems by building new objects from scratch,” as the video fades from images of prehistoric hunters to the Golden Gate bridge and 3D printers. The catch phrase for the institute, “we aim to design a new world of proteins,”¹⁴ clearly displays the change in scientific aims from naturalistic description and explanation of the natural world to *de novo* design. David Baker stated, “There’s a lot of things that nature has come up with just by randomly bumbling around” and “as we understand more and more of the basic principles, we ought to be able to do far better” [Zimmer 2017, n.p.], or along the same line the webpage of the Institute reads: “Evolution has done an incredible job making these proteins with exquisite functions, but nature can only do so much. Modern day challenges demand new proteins with new functions on a much faster timescale than evolution is capable of.”¹⁵ In addition to fulfilling the desired task (digest gluten, protect from flu, bind an opioid), the designed proteins that have been experimentally obtained are indeed in general superior in performance to the natural ones in terms of stability (“do far better”); natural protein folds are in general marginally stable and easily denatured, while several designed proteins were found to be highly stable and to withstand heat and chemical denaturation. The protein designers insist on the fact that these novel objects are designed “from scratch” stressing the anti-genealogic move by which these engineers wish to emancipate from (natural) historical constraints.

Here biology has led itself into its own negation, its Hegelian *Aufhebung*. Natural processes are outdated, obsolete, outcompeted. Interest in organic machines (Deleuze and Guattari might say *machines désirantes* (desiring machines) [Deleuze and Guattari 1972]) is replaced by creation of desired machines (*machines désirées*). Natural proteins are informed, and this information derives directly from evolutionary processes, from

developments in this area correspond to the introduction of the AlphaFold AI program based on Google’s DeepMind.

¹⁴ *What is the Institute for Protein Design?* (<https://www.youtube.com/watch?v=jBNCj2vqau0>, accessed November 26, 2019.)

¹⁵ <http://www.ipd.uw.edu/research/basic-areas/> (accessed April 4, 2018).

lived time. Proteins are active because they have acted before, their sequences have been shaped by time and action, and thus this information is encapsulated in the sequences of the corresponding genes. De novo designed proteins are also informed, but information is mind-derived and computer-generated. As such they can also be envisioned as an indirect derivation of evolutionary processes and their sequences also reflect the parameters of the energy functions that were used to build them, that is, human knowledge. The current blur in the nature/nurture divide invites us to consider these newly built second-order biological objects as part of the transition from biosphere to noosphere predicted by Vladimir Vernadsky (1945).

6 Intrinsic Disorder, Proteins at the Edge of Chaos

So far, we have treated the relation of sequence to fold and function within Anfinsen's principle that proteins have a minimum energy state and spontaneously fold to attain this shape, which in turn defines their function. But it is now known that in a number of cases proteins do not fold into a determined structural state but have a chaotic behavior and remain disordered [Habchi et al. 2014]. Instead of having a unique energy minimum they display multiple quasi-equivalent energetic states (Fig. 2). Initially, the existence of intrinsically disordered proteins (IDPs) was considered with much skepticism as this concept of natively unfolded proteins (as they are also called) contradicted the general paradigm of the structure to function relation that governed the understanding of structural biology. Also, their chaotic behavior rendered these objects recalcitrant to classical experimental structural description; this intractable character made it more comfortable to just disregard them for some time. It is now widely accepted that IDPs are common, and present in all forms of life. IDPs exist as a dynamic, permanently interconverting ensemble of structural states in which the protein chain is extended rather than collapsed into a globule-like state; the collection of the different states has been termed "a dancing protein cloud" by Vladimir Uversky (2016, p. 6681), a pioneer in the field. The order in 'normal' folded proteins is encoded in their sequences and naturally, so is also the disorder encoded in the sequences of these proteins (proteins at the edge of chaos as Uversky terms them). Disordered protein sequences display a bias in amino acid composition and are therefore also termed low complexity sequences. These are characterized by a depletion in order-promoting residues (for instance aromatic and hydrophobic residues) and enrichment in disorder-promoting residues (like glycine or polar residues). IDPs typically evolve faster than folded proteins as their sequences are less constrained; exact sequence matters less as long as a form of disorder is maintained. This structural plasticity transcends the classical one protein-one structure-one function concept and entails such proteins with functional polyvalence. The semantic catalogue associated with the description of this type of proteins includes terms as malleable machines, morphs, chameleon,

dual-personality, fuzzy, moonlighting, rheomorphic, or promiscuous sequences, all intended to convey this impression of plasticity. For instance, such proteins can interact with many different cellular partners and in an induced folding mechanism acquire a partial fold when interacting with a given binding partner. In this case, the folding code of the protein is not entirely intrinsically contained within the protein sequence but partially externalized in the sequence and fold of its cognate binding partner. Some IDPs, however, perform their function while remaining fully in a disordered state and have been termed stochastic machines [Xue et al. 2013].¹⁶

As the biological importance of disorder in protein function continues to be further documented, the one protein–one structure–one function concept needs to be gradually replaced or complemented by a more complex one protein–many structures–many functions paradigm. There appears to be an interesting trend of “blue-collar/white-collar” tasks distribution in cellular activities, folded proteins ensuring primary metabolism and transport functions while IDPs are more often involved in signaling and regulatory functions (specifically because IDPs do not adopt a stable 3D structure, they do not perform enzymatic functions which are based on the precise stereochemical positioning of catalytic chemical groups in the active sites; in contrast, IDPs are particularly suited for molecular recognition tasks involving multiple binding interactions). In that sense, folded proteins act while non-folders preferentially interact [Tompa et al. 2015]. Also, apparently there is a positive correlation between organismal complexity and prevalence of IDPs, the more complex the organism the more frequent IDPs are in its proteome [Habchi et al. 2014]. One should measure how much unrest the introduction of these anti-platonic objects introduced in the paradigm of a firm relation of form to function in biological entities.

7 Protein Fluctuation and Function

We have contrasted here the behavior of ‘normal’ ordered proteins to IDPs. This dichotomy is largely unjustified, and in fact some, as perhaps most prominently the biological physicist Hans Frauenfelder, have explained that proteins move. Even “normal” folded proteins are intrinsically dynamic and change shape on timescales ranging from the picosecond to the second (“A protein does not exist in a unique conformation but can assume a very large number of somewhat different conformations

¹⁶ “The [protein] complex works by random movements of a ‘stochastic machine’ not by cooperative conformational changes. Unlike typical machines, the different parts of the device are loosely connected, with random movements bringing components together” [Xue 2013, p. 1588]. In that sense, the word “loosely” is used in the same sense as in mathematics and probability theory as relating to a random process (here random motions).

or conformational substates” [Frauenfelder, McMahon, and Fenimore 2003, p. 8616; Henzler-Wildman and Kern 2007]. Some deeply rooted platonistic tendency fueled by the beautiful – static – images of protein structures produced by structural biology leads to the repeated oblivion of this fact.¹⁷ What is important here is not so much that proteins move, but that protein structural fluctuation is a condition of function. As Hans Frauenfelder puts it, “if a protein had just a single conformation, it could not function and would be dead like a stone” [cited in Kovac 2006, p. 564]. Structural fluctuation is critical for many aspects of protein function including signal transduction, enzymatic catalysis, ligand binding, and assembly of supramolecular complexes. In the end, the structure-function paradigm extends into a disorder-function concept.¹⁸

The concept of allostery is essential for the understanding of the relation of conformational change to function. The term “allostery” was coined by Jacques Monod [Monod, Changeux, and Jacob 1963] and denotes the ability of a protein to change its conformation in response to the binding of a ligand (Fig. 2). Initially framed in the context of multimeric enzymes whose activity is modulated by the binding of a ligand at a site distinct from the substrate binding site, this allostery concept has evolved considerably from a two-state to a multi-substate model and from multisubunit proteins to all types of proteins [Liu and Nussinov 2016]. This ability of proteins to change between substates in response to changes in their environment is considered so fundamental that it has been elevated to the level of a “second secret of life,”¹⁹ second only to the genetic code [Fenton 2008, p. 420].²⁰

17 Thanks to a series of technological innovations – such as solution X-ray scattering – structural biology becomes more and more apt to describe protein structural fluctuation [Palamini, Canciani, and Forneris 2016]. The discipline undergoes a kind of photography-to-cinematography transition. Novel graphic methods to display structural substates are being developed [Melvin and Salsbury 2016] and will perhaps play a heuristic role in the increased awareness of protein plasticity.

18 Aside from internal movement associated with transition from one substate to another, recent experiments enabled by development of specific spectroscopic methods have revealed that movement (in the form of displacement of the object in space) is also associated to enzymatic catalysis per se. Enzymes were found to be propelled, to leap by ballistic-like movements in a length scale of 100 nm when performing their catalytic function. Such leaps are more frequent at high substrate concentration, so that enzymes collectively distribute in a gradient inverse to the concentration gradient of the substrate, in a process that was designated antichemotaxis [Jee et al. 2018].

19 Agnès Ullmann provides a recollection of the discovery of the allostery concept: “By the end of 1961, one evening quite late Jacques Monod walked into my laboratory looking rather tired and worried. Monod stood silently at my bench and after a few long minutes he said: ‘*I think I have discovered the second secret of life.*’ I was quite alarmed by this unexpected revelation and asked him if he needed a glass of whisky. After the second or maybe the third glass, he explained the discovery, which he had already given a name: ‘allostery.’ ” [Ullmann 2011, p. 1029].

20 In a following section, we will question this order and suggest that the possibility exists that allosteric behavior has a temporal primacy and that – if at all – it is the genetic code that might be qualified as ‘second secret of life.’

To describe the existence of structural substates in proteins, for instance determined by allostery, the energy landscape model is used (we have already mentioned this model in the context of protein folding) [Henzler-Wildman and Kern 2007]. While the energy landscape is generally represented as a funnel-shaped landscape with a single minimum, protein landscapes are better understood as rugged landscapes with a number of local minima that can be explored. These local energy minima are separated by barriers that can be overcome thanks to heat and only become effective at very low temperatures (which in a way ‘freeze’ protein motion); higher energy barriers are overcome by the binding of a ligand or various changes in the solvent condition or physical conditions. Frauenfelder states poetically that “to perform functions, proteins must dance in the energy landscape” [Frauenfelder 2017, p. 1]. It should be noted that there is no fundamental contradiction between Anfinsen’s principle and the existence of protein conformation substates as Anfinsen considers energy “of the whole system,” protein and solvent, stressing that protein folds are fundamentally co-constructions of the protein sequence and the environment.

8 Protein Coalescence, Flocks of Proteins

Gradually over the last few years, a remarkable new view of the collective behavior of proteins in living cells has emerged. While proteins in general act in a soluble monomeric state or as multimeric supramolecular complexes of defined composition, they also display a collective behavior involving coordinated self-associations leading to phase transitions [Shin and Brangwynne 2017; Kaganovich 2017]. In response to environmental changes, a large number of proteins undergo a transition from a soluble to a coalescent state which can correspond to different physical states ranging from liquid and liquid crystal-like to solid (Fig. 2). This phenomenon – which has been designated under a variety of terms of which protein coalescence is perhaps the most satisfactory – encompasses a variety of biological processes that have been so far considered in isolation and now receive a more unified consideration because all involve new properties conferred by the collective behavior of proteins [Saarikangas and Barral 2016]. At the molecular level, these protein coalescence processes involve often but not always protein domains that tend to be intrinsically disordered as described here above. The structurally dynamic character of the IDPs or protein domains appears to be associated with their ability to form such coalescent states. Such phase transitions can correspond to the formation of so-called membrane-less organelles in which the coalescence of proteins (and often associated RNAs) leads to the formation of cellular compartments which are not bound by a lipid membrane. These protein coalescence structures (which receive different names like granules, droplets or protein bodies) often form in response to an environmental change (typically a stress, like nutrient depletion or heat). Some exist only transiently and

dissolve when the environmental signal subsides, others become permanent and even, as we shall see, can be transmitted to the cell progeny as genetic material. Coalescent structures can be composed of up to several hundreds of different protein (and RNA) types or result from the collective behavior of a single protein type. Protein coalescence processes serve different cellular functions including sensing, filtering, protective storage, protein inactivation, amplification of signal or enzymatic activity and memory. While protein coalescence is often an adaptive phenomenon, collective protein behavior also leads to pathological situations, such as neurodegenerative protein deposition diseases like Parkinson's, Alzheimer's, and Huntington's diseases, which typically result from the collective behavior and coalescence of proteins. Physically, these protein flocks can correspond to liquid droplets and involve transient interactions between largely unstructured proteins, but they can also be highly ordered or disordered solids that assemble and disassemble at various rates. While the exact relation between the different condensed states is often unclear, liquid condensates can further mature into liquid crystal-like and solid states. When an ordered solid state is involved, it most often corresponds to a protein state termed amyloid, an ordered arrangement of protein subunits into a fibrillar polymer with a cross- β structure formed by the stacking of β -strand structure elements perpendicular to the fibril axis.

An interesting classification of these protein coalescence phenomena based not so much on the physical state of the protein condensate but on its biological properties has been proposed [Saarikangas and Barral 2016]. This classification distinguishes *adaptopods* that form in response to a signal but dissolve when the inducing signal is removed from *mnemons* which are signal-induced but self-maintained (remain after the signal is removed). Mnemons thus fix a new state irreversibly. The Whi3 protein found in yeast provides such an example of proteins that – by collective behavior – act as a device capable of encoding the memory of past events. In yeast cells, exposure to the pheromone from the opposite mating-type induces an arrest of the cell cycle; the cell stops dividing and differentiates a mating structure termed *shmoo* (because it resembles the cartoon strip creature created by Al Capp). If, however, no mating partner is reached within a reasonable time, the cell cycle resumes and cells become refractory to a new pheromone exposure (they will not be fooled again into growth arrest and shmoo-formation). This acquired behavior results from the pheromone-induced coalescence of the Whi3 proteins which thus “encodes memory of deceptive encounters” [Caudron and Barral 2014, p. 100]. This mechanism of cellular memory based on collective protein behavior was reported to also occur in long-term memory in invertebrates and mammals. In yeast, mnemons are defined by the fact that they are not inherited in cell divisions. The Whi3 mnemon is physically retained in the mother cell that permanently becomes refractory to pheromone stimulation, but the daughter cells do not inherit memory of the event and are

competent for mating.²¹ This distinction in heritability of the protein coalescent structure defines a last category of protein assemblies termed *prions* that are transmitted to cell progeny and thus represent *genetic* elements that will be described in the following section.

These examples illustrate that in addition to their individual activities, collective behavior of proteins in the form of phase transition in a variety of modes entails emergent properties that are developed in a multitude of cellular processes ranging from adaptation to memory and inheritance.

9 Inheritable Collective Folds, Proteins as Genes

Prions illustrate a specific case of phase transition, in which the coalescent state becomes transmissible. Prions are infectious proteins, they propagate their fold. The term “prion” was coined in the context of the study of fatal neurodegenerative diseases in different mammals, such as Creutzfeldt-Jacob disease and Kuru in humans and Mad Cow disease in cattle [Colby and Prusiner 2011]. But there are many other prion proteins notably in fungi where they were initially identified as non-Mendelian genetic elements [Harvey, Chen, and Jarosz 2018]. These fungal prion proteins are not pathological; they have adaptive, functional roles in cells. For the mammalian PrP prion, the prion state corresponds to a molecular mishap, a *misfolding*. Prion diseases, like other neurodegenerative diseases such as Parkinson’s and Alzheimer’s, are misfolding diseases. In contrast, in the case of fungal prions, the prion fold is the native fold. Importantly, prion folding can only be achieved cooperatively and not by an isolated molecule. The information contained in the DNA sequence can only be actualized when several protein monomers interact to collectively attain the prion fold as an assembly of molecules. They undergo a phase transition of the kind described in the above paragraph into a solid and highly ordered fibrillar amyloid aggregate. Once formed, this fold becomes self-propagating, in the sense that it directs the folding of other protein units to the same prion fold. The prion fold directly transmits this structural information by contact and imposes it on other molecules of the same kind. The prion template grows into a fibril by incorporation of monomers and ultimately breaks down to smaller fragments, the extremities of which then act as new templates for transmission of the fold, thus ensuring replication of the propagating entity.

For reasons of personal taste and because it is the first prion whose high-resolution structure has been solved, we will turn to the specific example of the s (or [Het-s]) prion of the filamentous fungus *Podospora anserina* to illustrate the

²¹ In the yeast-jargon, the daughter cell buds off by mitosis from the mother cell, so formally the relation of the two is in no way genetically analogous to the mother–daughter relation in humans.

prion principle [Wasmer et al. 2008]. Like a machine connected to other machines in various configurations, this protein stands at the crossroads of several biological phenomena and depending on the perspective that is adopted can be viewed as an inactive mutant, a protein defining biological individuality or an ultra-selfish genetic element. A description of the complex natural history of this protein can be found elsewhere [Daskalov and Saupe 2015]. It might be sufficient to state here that fungi expressing this protein can display two alternate phenotypes termed *s* and *s** that differ in their reactivity towards a genetically distinct strain termed *S*. *s** strains are able to fuse with *S* (*s** and *S* are compatible), while *s* strains cannot and produce a cell death reaction when in contact with *S* (*s* and *S* are incompatible). This difference in phenotype is not defined by a difference in DNA sequence but results from an epigenetic difference, residing only in the folding state of the *s* protein. It is important to understand that this fold contains transmissible information, hence genetic information. This information is transmitted horizontally from cell to cell and inherited vertically in cell divisions and in sexual reproduction. Experiments analogous to the classical work of Oswald Avery, Colin McLeod and Maclyn McCarty that established DNA as the ‘transforming principle’ (the genetic molecule) have been carried out with the *s* prion protein and amount to defining the protein as genetic material [Fink 2005]. Formally, the fold – derived from the collective behavior of these proteins – is a gene.

In most proteins, the folding is delayed typically by microseconds to seconds, but here, with this prion fold, the fold can be considerably delayed. The prion folding can in fact be deferred indefinitely (or at least never be actualized in the life span of the organism). The proteins in the unfolded and folded states are identical in terms of chemical composition, they are the same and yet they differ and this difference is the fold; in one state information is actualized, in the other not. This fold, this information now actualized is self-propagating. The prion protein affects itself, but the other itself, its *deferred* self (prion conversion is an auto-affection but also a hetero-affection). This biological system, with some of its specificities, bends and stretches the vocabulary and nomenclature of modern biology to its limits by making the genotype/phenotype divide undecidable.²²

²² The textual analogy in molecular biology is a classical one and perhaps the *différance* proposed by Jacques Derrida in the linguistic context can be of help to describe the ambiguous signifier-to-signified relation in this system. Derrida gradually developed *différance* to describe the process by which a language (or any system of code) historically constitutes itself as a system of differences, producing spacing between elements that get defined by difference in relation to each other. The term carries a radical polysemy as it contains altogether the notions of temporal delay (*différé*), of alterity of dissemblance (*différence*), of allergic or polemic alterity (*différend*) and finally the notion of production of difference (*-ance* suffix). The information specifying the protein fold is present in the DNA of the corresponding gene (the sequence makes sign for the protein). But the relation of signifier-to-signified in the sign is deferred. This sign associates a single-but-multiple signified (unfolded non-prion/folded prion states of the *same* protein) to a given signifier (sequence). The

10 In the Beginning There Was the No Verb, Proteins, and the Origin of Life

Several aspects in the understanding of the collective behavior of proteins lead to some interesting ideas regarding the possible origin of life (however, it might be defined). Very basic requirements for cellular life are some form of templating underlying heritability and self-perpetuation of structure and some form of confinement with at least partial definition of an inside and an outside, the roles of template and boundary being attributed classically and respectively to nucleic acids (DNA and/or RNA) and the lipid bilayer membrane. Some researchers currently contemplate the hypothesis that proteins in the form of amyloid condensates might represent such initial self-templating molecules [Maury 2015]. It has been possible to obtain spontaneous formation and replication of amyloid peptides in Miller soup pre-biotic conditions [Greenwald, Friedman, and Riek 2016; Rout et al. 2018]. In these experiments, peptides are extended by amino acid addition selected by direct structural templating, not through decoding of a triplet sequence. The same is true for certain types of α -helical peptides, raising the possibility that direct templating in protein synthesis predates development of the genetic code. Also, in the frame of this view, amyloid (collective) folds would represent ancient, primitive folds and subsequent evolution of monomeric folded proteins a further step. Formation of complex folds would thus represent emancipation from this primitive rather monotonous state. Complex monomeric folding requires generally a certain length of the protein chain (a folded domain generally requires about 20–30 amino acids in length to fold). The increase in length of the protein chain might represent a later evolutionary stage compared to the small length peptide condensates that could form in prebiotic conditions. It is interesting to consider that this path might be taken in reverse during age-dependent

sequence signifies the prion fold, but the fold is deferred. In an anticipation of the last section of this text, we note that the temporal delay component of the *différance* is presented as a delay of the realization of desire (“*Différer* in this sense is to temporize, to take recourse, consciously or unconsciously, in the temporal and temporizing mediation in a detour that suspends the accomplishment or fulfillment of ‘desire’” [Derrida 1982, p. 8] (“*Différer* en ce sens, c’est temporiser, c’est recourir, consciemment ou inconsciemment, à la médiation temporelle et temporisatrice d’un détour suspendant l’accomplissement ou le remplissement du ‘désir’” [Derrida 1972, p. 8])). The fold is deferred, this difference is historically produced and by external reference (as a co-construction with the environment), there is *différance* in and of the fold. This prion system provides illustrations of how alterity can emerge, how to produce difference from the same [Daskalov and Saupé 2015]. In this system, an alterity of dissemblance (in sequence or structure) is differentiated into (produces) an alterity of allergy (a conflict, incompatibility). The dualistic oppositions in this system (*s/s** and *s/S*) are only decisive by external reference. There is *différance* in as much as the fold is delayed and in as much as it is productive; it is what actualizes the biological alterity between *s* and *S*. In passing, just like the ‘a’ in Derrida’s *différance* is inaudible, the *s/S* difference is graphical and inaudible unless it is outspoken.

protein misfolding, in which proteins due to deficiency in protein quality control mechanisms or chemical alteration form amyloid aggregates at the expense of the normal folded states. It is Ernst Haeckel's ontogeny-recapitulates-phylogeny in reverse [Haeckel 1866].

The protein-based phase separation processes, mentioned in a previous section, offer a form of macromolecular organization that separates an interior from an exterior milieu and just as some cell organelles – which are considered as remnants of endosymbiotic cells – are bound by lipid bilayers, others are not and it is now envisioned that phase-separated protein bodies might have played a role as protocells at the origin of life [Stroberg and Schnell 2017]. Phase-separated droplets are capable of filtering out or concentrating certain chemicals and ions and are capable of mitosis-like division processes. In this form of individualization, there is formation of two compartments but no actual distinct material frontier separating them (as in the case of the lipid bilayer). It is the intrinsic interaction of the phase-separated entities that provides the compartmentation function.

While these views of the origin of life remain obviously speculative to the highest degree, they provide an alternative to the perspective strongly centered on the main informational molecule in extant organism, DNA, and to draw on Lily Kay's use of the biblical "in the beginning was the verb" metaphor framing the importance of information theory in the historical development of molecular biology [Kay 2000], maybe one could say that the verb was not made flesh but was the flesh right from the start. As an article on the origin of the word "protein" explains, "The name 'protein' [was] derived from the Greek adjective πρωτεῖος, meaning of the first rank or position" [Vickery 1950, p. 387]. The etymology relates to the fact that the described substance was supposed to be a primary stuff for animal nutrition produced by plants, so in fact does not imply temporal primacy, but Valère Novarina proposes that in some instance erroneous etymologies might be more effective than real ones [Novarina 2017, p. 22].

11 Proteins as Monads and Desiring Machines

The all-encompassing materialistic philosophy of Gilles Deleuze and Félix Guattari was defined by some as a biophilosophy and analyzed for instance as containing a critique of the neo-Darwinian modern synthesis [Marks 2006]. But aside and maybe complementary to this aspect, what is highly original in Deleuze's and Guattari's reception of molecular biology is their attention to protein activity and folding. In an anti-genealogic gesture, they find more interest in the actual living machine than in its genealogical origin. There are explicit references in both Deleuze's *Le Pli (The Fold)* and Deleuze's and Guattari's *Mille Plateaux (A Thousand Plateaus)* and *L'Anti-Œdipe (Anti-Oedipus)* to protein folding. In *Le Pli*, Deleuze proposes, following Leibniz, a description of organic development based on folding/unfolding processes and places

protein folding at the very root of such organic developmental processes [Deleuze 1988, p. 15]. Deleuze subscribes to the Leibnizian machinistic view of organic life: “But the machines of nature, namely, living bodies, are still machines in their smallest parts *ad infinitum*” [Leibniz 1898, p. 254],²³ and in this way he distinguishes *machinism* from *mechanism*: “A mechanism is faulty not for being too artificial to account for living matter, but for not being mechanical enough, for not being adequately machined” [Deleuze 2006, p. 8].²⁴ Protein complexes with their many folds and motions conform to the baroque aesthetics as defined by Deleuze in his essay on Leibniz and the Baroque [Deleuze 1988, p. 6] and conform to Deleuze’s definition of organic machines as baroque machines [Deleuze 1988, p. 12] (Fig. 1). Following Deleuze in his reading of Leibniz and putting aside the theological dimension of Leibniz’ *La Monadologie*, it might be possible in some aspects to qualify proteins as *monads*.²⁵ While of course, in chemical and physical terms, proteins do not represent “simple substances” that cannot be further subdivided, it might be argued that any further subdivision below the level of the individual protein cannot be achieved without losing biological meaning, a fragment of the protein does not make sense any more, it is not made of machines but of (chemical) parts. As such, proteins might qualify as the Leibnizian “Atoms of Nature” [Leibniz 1991, p. 125]. From these elementary machines, by a series of higher order organization steps, macromolecular complexes, organelles, cells, organs and organisms can be assembled. Deleuze stresses that this organization, this formation of organisms, can only be understood if one assumes that the elementary components are already folded, organized.²⁶ Another relevant property of monads is a certain level of perfection and autonomy (“they have in them a certain perfection [...];

23 “Mais les machines de la nature, c’est à dire les corps vivants, sont encore des machines dans leurs moindres parties, jusqu’à l’infini” [Leibniz 1991, p. 162].

24 “Le tort du mécanisme ce n’est pas d’être trop artificiel pour rendre compte du vivant, mais de ne pas l’être assez, de ne pas être assez machine” [Deleuze 1988, p. 12].

25 “The Monad, of which we shall here speak, is nothing but a *simple* substance, which enters into compounds. By ‘simple’ is meant ‘without parts’” [Leibniz 1898, p. 217]. (“La Monade, dont nous parlerons ici, n’est autre chose, qu’une substance *simple* qui entre dans les composés; simple, c’est-à-dire, sans parties” [Leibniz 1991, pp. 123–124].)

26 “On the other hand, the formation of the organism would remain an improbable mystery, or a miracle, even if matter were to divide infinitely into independent points. But it becomes increasingly probable and natural when an infinity of indeterminate states is given (already folded over each other), each of which includes a cohesion at its level, somewhat like the improbability of forming a word by chance with separate letters, but with far more likelihood with syllables or inflections” [Deleuze 2006, p. 7]. (“D’autre part, la formation de l’organisme resterait un mystère improbable ou un miracle si la matière se divisait même à l’infini en points indépendants, mais devient de plus en plus probable et naturelle quand on se donne une infinité d’états intermédiaires (déjà repliés) dont chacun comporte une cohésion, à son niveau, un peu comme il est improbable de former au hasard un mot avec des lettres séparées, mais beaucoup plus probable avec des syllables ou des flexions” [Deleuze 1988, p. 10].)

they have a certain self-sufficiency [...] which makes them the sources of their internal activities” [Leibniz 1898, p. 229]),²⁷ which is in line with the ability to fold autonomously into a unique fold associated with function. They are perfect in the sense that they attain an energy minimum. Then importantly monads have *perceptions* and *appétites*.²⁸ The concept of allostery by which a protein perceives and changes in response to ligands (or physical conditions) illustrates that sentience, understood as “the capacity to exhibit a variety of potential internal states, which respond to the immediate state of the environment” [Kovac 2006, p. 564], extends *cognition* to the level of proteins (at the individual protein level, at the level of the higher order macromolecular complexes or the collective phase-transition behaviors). Jacques Monod – one of the main sources of reading on molecular biology for Deleuze and Guattari together with François Jacob’s *The Logic of Life (La Logique du Vivant, 1970)* – was apparently the first to use the term “cognition” for substrate discrimination and enzyme function. Deleuze describes chemical affinities of enzymes as “perceptions and elections” and plays with the temptation to term these “souls” [Deleuze 1981]. Ladislav Kovac, in a remarkable text entitled “Life, chemistry and cognition,” frames the contours of what he terms “cognitive biology,” a concept centered on the notion that cognition is ultimately and fundamentally embodied [Kovac 2006, p. 563]. Kovac insists that cognition is a dual process of sensation and action and convincingly installs the proteins at the very basis of a hierarchical, multilayer, nested sentience. The central concept allowing this articulation between sensation and action is allostery, that is, the ability of a protein to undergo a conformational change in response to ligand binding (or in a broader sense to other physical cues such as light, heat, or mechanical forces) and to display changed functional properties in response to this signal.²⁹ Current biotechnological projects exploit the “cognitive ability” of proteins (and/or nucleic acids) to construct various Boolean logic gates in order to create biocomputing nanorobots that can be interfaced with living biological systems [Katz 2015].

Furthermore, a monad is defined as a reflection of the entire universe taken from a certain perspective (“as every Monad is, in its own way, a mirror of the universe” [Leibniz 1898, p. 253]).³⁰ Richard Lewontin states that organisms are molded

27 “elles ont en elles une certaine perfection [...], il y a une suffisance [...] qui les rend sources de leurs actions internes” [Leibniz 1991, pp. 133–134].

28 “*Perception*” and “*Appétition*” [Leibniz 1991, pp. 129, 131].

29 “Recognition is followed by an action. A ligand is a signal. In contrast to standard chemical interactions, binding energy is not fully dissipated as heat, but is used partly for molecular work – a specific pre-programmed change in the conformation of the protein. In this way, the signal is transmitted from one site on the protein to another. The transmission takes place in four-dimensional space, as it involves time as a coordinate, and this process gives biochemistry its vectoriality. [...] It is appropriate to regard most protein molecules as molecular engines. Hence, molecular cognition consists of molecular sensation – which has two inseparable aspects, recognition and signification – and molecular action” [Kovac 2006, p. 565].

30 “toute Monade étant un miroir de l’univers à sa mode” [Leibniz 1991, p. 161].

by and are thus molds of their environments [Lewontin 2001]. Molecular evolution shapes protein sequences/folds in response to specific biotic and abiotic cues in the environment, as such proteins seen as monads reflect a certain point of view on the universe. Perhaps the case of ancestral protein reconstruction experiments can serve to illustrate this principle. Sequences of very ancient enzymes dating back as far as LUCA (Last Universal Common Ancestor) can be expressed, produced, and structurally characterized. By analyzing their properties, thermal stability, thermal optimum, pH dependence, and substrate specificity, one can devise information about the physico-chemical conditions on earth a billion years ago [Akanuma 2017]. A single protein reflects its *Umwelt* [Uexküll 1909], mirrors its own perspective on the universe.

In the two volumes of *Capitalisme et Schizophrénie (Capitalism and Schizophrenia)*, Deleuze and Guattari place the concept of *desiring machine* as the basis of desiring-production. Desiring machines are connected with each other, machines branched on the flux of another machine, these loose and changing connections between elementary machines build up organisms and at higher levels social structures [Andoka 2012]. In their development, a critical step occurs in Section 2 of Chapter 4 of *Anti-Oedipus*, “beyond vitalism and mechanism” [Deleuze and Guattari 2000]. Both vitalism and mechanism deprive the machine of a direct relation to desire.³¹ In mechanism, desire is an emergent effect of the abstracted structural unity explaining the organism. In vitalism, machines are subordinated to the desire of the individual unity of the organism. The biologist frames the understanding of life with the concept of function, a concept genuine to biology, says Peter Wolynes [Ferreiro, Komives, and Wolynes 2018], which might provisionally replace desire in the above sentences. Deleuze and Guattari aim at shattering both perspectives by putting into question both the individuality of the organism and the structural unity of the machine. What is achieved by dissolving the mechanistic and vitalistic perspectives is the direct

³¹ “From machines, mechanism abstracts a *structural unity* in terms of which it explains the functioning of the organism. Vitalism invokes an *individual and specific unity* of the living, which every machine presupposes insofar as it is subordinate to organic continuance, and insofar as it extends the latter’s autonomous formations on the outside. But it should be noted that, in one way or another, the machine and desire thus remain in an extrinsic relationship, either because desire appears as an effect determined by a system of mechanical causes, or because the machine is itself a system of means in terms of the aims of desire. The link between the two remains secondary and indirect, both in the new means appropriated by desire and in the derived desires produced by the machines” [Deleuze and Guattari 2000, p. 284]. (“Le mécanisme abstrait des machines une *unité structurale* d’après laquelle il explique le fonctionnement de l’organisme. Le vitalisme invoque une *unité individuelle et spécifique* du vivant, que toute machine suppose en tant qu’elle se subordonne à la persistance organique et en prolonge à l’extérieur les formations autonomes. Mais on remarquera que, d’une manière ou d’une autre, la machine et le désir restent ainsi dans un rapport extrinsèque, soit que le désir apparaisse comme un effet déterminé par un système de causes mécaniques, soit que la machine soit elle-même un système de moyens en fonction des fins du désir. Le lien reste secondaire et indirect entre les deux, aussi bien dans les nouveaux moyens que le désir s’approprie que dans les désirs dérivés que suscitent les machines” [Deleuze and Guattari 1972, p. 337].)

connection between the machine and desire.³² This may have been overlooked, but Deleuze and Guattari *directly* define proteins as the elementary desiring machines: “Proteins are both products and units of production; they are what constitutes the unconscious as a cycle or as the autoproduction of the unconscious – the ultimate molecular elements in the arrangement of the desiring-machines and the syntheses of desire” [Deleuze and Guattari 2000, p. 290].³³ Central to this point of view is the aforementioned concept of allostery that articulates recognition and action, converting the Leibnizian *perceptions* and *appetites* to desiring-production. Even the most trivial biotechnological application of protein function/desire (such as the use of lipases in washing powder) illustrates how the mechanistic or vitalistic unity can be undone and machine connections reshuffled and reorganized. An interesting mirror image of proteins seen as desiring machines can be found in the development of a chemical concept termed *frustration*. There is frustration in a protein sequence segment when fold stability conflicts with the plasticity required to perform various binding functions. Which regions of a protein sequence are frustrated (present a conflict between stability and plasticity) can be determined and Peter Wolynes proposes that analyzing protein frustration is a way to understand the goals of a protein, to “distinguish meaning from noise” and that “analyzing frustration is a tool to search for meaning,” as if finding where frustration is will help in defining desire [Ferreiro, Komives, and Wolynes 2018, p. 71].

We aimed here at discussing in which sense proteins can be viewed as active matter, in both their individual and collective behavior, Deleuze and Guattari might say on their molecular and molar side. In the process, the materialism developed by these post-structuralist philosophers rooted in the Leibnizian theory of “les petites perceptions” (“the small perceptions”) as Deleuze defines it [Deleuze 1981] represents a useful guide. Many problematic aspects of contemporary biology could be thought of in the terms developed in *L’Anti-Œdipe* and *Mille Plateaux*. Organismal unity in all phylogenetic branches is fundamentally reframed by metagenomics,

32 “Once the structural unity of the machine has been undone, once the personal and specific unity of the living has been laid to rest, a direct link is perceived between the machine and desire, the machine passes to the heart of desire, the machine is desiring and desire, machined. Desire is not in the subject, but the machine in desire – with the residual subject off to the side, alongside the machine, around the entire periphery, a parasite of machines, an accessory of vertebro-machinate desire” [Deleuze and Guattari 2000, p. 285]. (“Une fois défaite l’unité structurale de la machine, une fois déposée l’unité personnelle et spécifique du vivant, un lien direct apparaît entre la machine et le désir, la machine passe au cœur du désir, la machine est désirante et le désir, machiné. Ce n’est pas le désir qui est dans le sujet, mais la machine dans le désir – et le sujet résiduel est de l’autre côté, à côté de la machine, sur tout le pourtour, parasite des machines, accessoire du désir vertébro-machiné” [Deleuze and Guattari 1972, p. 339].)

33 “Les protéines sont à la fois produites et unités de production, c’est elles qui constituent l’inconscient comme cycle ou l’autoproduction de l’inconscient, ultimes éléments moléculaires dans l’agencement des machines désirantes et des synthèses du désir” [Deleuze and Guattari 1972, p. 345].

prevalence of horizontal transfer questions species genealogy, phylogenetic trees are replaced by reticulated (rhizome-like) trees [Lopez and Baptiste 2009], and exaptation has been documented down to the molecular level. The interesting aspect here is that in the late 1970s, at the time these works appeared, molecular biology offered a much simpler, cleaner, and ordered picture of the living world. Nevertheless, the reading of this picture by Deleuze and Guattari is much more in line with the multiple, deconstructed, reticulated image emerging from our twenty-first-century molecular biology.

Bibliography

- Akanuma S. Characterization of Reconstructed Ancestral Proteins Suggests a Change in Temperature of the Ancient Biosphere. *Life* 2017, 7 (3), 33, DOI: 10.3390/life7030033.
- Andoka F. Machine désirante et subjectivité dans l'Anti-Œdipe de Deleuze et Guattari. *Philosophique* 2012, 15, 85–94, DOI: 10.4000/philosophique.659.
- Andraut R. The Machine Analogy in Medicine: A Comparative Approach of Leibniz and his Contemporaries. In: Smith M, Nachtomy O, eds. *Machines of Nature and Corporeal Substances*. Dordrecht, Springer, 2011, 95–114.
- Anfinsen CB. Principles that Govern the Folding of Protein Chains. *Science* 1973, 181 (4096), 223–230, DOI: 10.1126/science.181.4096.223.
- Baldwin RL. Clash between Energy Landscape Theory and Foldon-Dependent Protein Folding. *PNAS* 2017, 114 (32), 8442–8443, DOI: 10.1073/pnas.1709133114.
- Bergson H. *L'Évolution Créatrice*. Paris, PUF, 2013 [1907].
- Bokori-Brown M, Martin TG, Naylor CE, Basak AK, Titball RW, Savva CG. Cryo-EM Structure of Lysin Pore Elucidates Membrane Insertion by an Aerolysin Family Protein. *Nature Communications* 2016, 7, 11293, DOI: 10.1038/ncomms11293.
- Butterfield GL, Lajoie MJ, Gustafson HH, et al. Evolution of a Designed Protein Assembly Encapsulating its own RNA Genome. *Nature* 2017, 552 (7685), 415–420, DOI: 10.1038/nature25157.
- Caudron F, Barral Y. Mnemons: Encoding Memory by Protein Super-Assembly. *Microbial Cell* 2014, 1 (3), 100–102, DOI: 10.15698/mic2014.01.134.
- Colby DW, Prusiner SB. Prions. *Cold Spring Harbor Perspectives in Biology* 2011, 3 (1), a006833, DOI: 10.1101/cshperspect.a006833.
- Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popović Z, Players F. Predicting Protein Structures with a Multiplayer Online Game. *Nature* 2010, 466 (7307), 756–760, DOI: 10.1038/nature09304.
- Daskalov A, Saupe SJ. As a Toxin Dies a Prion Comes to Life: A Tentative Natural History of the [Het-s] Prion. *Prion* 2015, 9 (3), 184–189.
- Deleuze G. *Le Pli. Leibniz et le Baroque*. Paris, Les Editions de Minuit, 1988.
- Deleuze G. *The Fold. Leibniz and the Baroque*. London, Continuum, 2006.
- Deleuze G. *Cours sur Spinoza du 6 Janvier 1981, 2^{de} Partie*, Transcription: Véronique Delannay et Jean-Charles Jarrell, Université Paris 8. (Accessed June 11, 2020, at http://www2.univ-paris8.fr/deleuze/article.php?id_article=9).
- Deleuze G, Guattari F. *L'Anti-Oedipe*. Paris, Les Editions de Minuit, 1972.
- Deleuze G, Guattari F. *Mille Plateaux*. Paris, Les Editions de Minuit, 1980.
- Deleuze G, Guattari F. *Anti-Oedipus*. Minneapolis, University of Minnesota Press, 2000.

- Deleuze G, Guattari F. *A Thousand Plateaus*. Minneapolis, University of Minnesota Press, 2005.
- Derrida J. *Marges de la Philosophie*. Paris, Les Editions de Minuit, 1972.
- Derrida J. *Margins of Philosophy*. Chicago, The University of Chicago Press, 1982.
- Dill KA, MacCallum JL. The Protein-Folding Problem, 50 years on. *Science* 2012, 338 (6110), 1042–1046, DOI:10.1126/science.1219021.
- Eaton WA, Wolynes PG. Theory, Simulations, and Experiments Show that Proteins Fold by Multiple Pathways. *PNAS* 2017, 114 (46), E9759–E9760, DOI: 10.1073/pnas.1716444114.
- Englander SW, Mayne L. The Case for Defined Protein Folding Pathways. *PNAS* 2017, 114 (31), 8253–8258, DOI: 10.1073/pnas.1706196114.
- Fenton AW. Allostery: An Illustrated Definition for the ‘Second Secret of Life’. *Trends in Biochemical Sciences*. 2008, 33 (9), 420–425, DOI: 10.1016/j.tibs.2008.05.009.
- Ferreiro DU, Komives EA, Wolynes PG. Frustration, Function and Folding. *Current Opinion in Structural Biology* 2018, 48, 68–73, DOI: 10.1016/j.sbi.2017.09.006.
- Fink GR. A Transforming Principle. *Cell* 2005, 120 (2), 153–154, DOI: 10.1016/j.cell.2005.01.004.
- Frauenfelder H. Physical Concepts in Biology-Past and Future. *Physical Biology* 2017, 14 (1), 010402, DOI: 10.1088/1478-3975/aa559f.
- Frauenfelder H, McMahon BH, Fenimore PW. Myoglobin: The Hydrogen Atom of Biology and a Paradigm of Complexity. *PNAS* 2003, 100 (15), 8615–8617, DOI: 10.1073/pnas.1633688100.
- Greenwald J, Friedmann MP, Riek R. Amyloid Aggregates Arise from Amino Acid Condensations under Prebiotic Conditions. *Angewandte Chemie Internationale Edition* 2016, 55 (38), 11609–11613, DOI: 10.1002/anie.201605321.
- Guo P, Noji H, Yengo CM, Zhao Z, Grainge I. Biological Nanomotors with a Revolution, Linear, or Rotation Motion Mechanism. *Microbiology and Molecular Biology Reviews* 2016, 80 (1), 161–186, DOI: 10.1128/mmb.00056-15.
- Habchi J, Tompa P, Longhi S, Uversky VN. Introducing Protein Intrinsic Disorder. *Chemical Reviews* 2014, 114 (13), 6561–6588, DOI:10.1021/cr400514h.
- Haeckel E. *Generelle Morphologie der Organismen*. Berlin, Georg Reimer, 1866.
- Harvey ZH, Chen Y, Jarosz DF. Protein-Based Inheritance: Epigenetics beyond the Chromosome. *Molecular Cell* 2018, 69 (2), 195–202, DOI: 10.1016/j.molcel.2017.10.030.
- Henzler-Wildman K, Kern D. Dynamic Personalities of Proteins. *Nature* 2007, 450 (7172), 964–972, DOI: 10.1038/nature06522.
- Huang PS, Boyken SE, Baker D. The Coming of Age of de Novo Protein Design. *Nature* 2016, 537 (7620), 320–327, DOI: 10.1038/nature19946.
- Ionov L. Polymeric Actuators. *Langmuir* 2014, 31 (18), 5015–5024, DOI: 10.1021/la503407z.
- Jackson SE, Suma A, Micheletti C. How to Fold Intricately: Using Theory and Experiments to Unravel the Properties of Knotted Proteins. *Current Opinion in Structural Biology* 2017, 42, 6–14, DOI: 10.1016/j.sbi.2016.10.002.
- Jee AY, Dutta S, Cho YK, Tlusty T, Granick S. Enzyme Leaps Fuel Antichemotaxis. *PNAS* 2018, 115 (1), 14–18, DOI: 10.1073/Pnas.1717844115.
- Judson HF. *Max Delbruck. The Eighth Day of Creation*. New York, Cold Spring Harbor Press, 1996.
- Kaganovich D. There Is an Inclusion for That: Material Properties of Protein Granules Provide a Platform for Building Diverse Cellular Functions. *Trends in Biochemical Sciences* 2017, 42 (10), 765–776, DOI: 10.1016/j.tibs.2017.08.002.
- Katz E. Biocomputing – Tools, Aims, Perspectives. *Curr Opin Biotechnol* 2015, 34, 202–208.
- Kay LE. *Who Wrote the Book of Life? A History of the Genetic Code*. Stanford, Stanford University Press, 2000.
- Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* 1958, 181 (4610), 662–666, DOI: 10.1038/181662a0.

- Kováč L. Life, Chemistry and Cognition: Conceiving Life as Knowledge Embodied in Sentient Chemical Systems Might Provide New Insights into the Nature of Cognition. *EMBO Reports* 2006, 7 (6), 562–566, DOI: 10.1038/sj.embor.7400717.
- Kryshtafovych A, Monastyrskyy B, Fidelis K, Moulton J, Schwede T, Tramontano A. Evaluation of the Template-Based Modeling in CASP12. *Proteins* 2018, 86 (Suppl. 1), 321–334, DOI: 10.1002/prot.25425.
- Laskey RA, Honda BM, Mills AD, Finch JT. Nucleosomes are Assembled by an Acidic Protein which Binds Histones and Transfers them to DNA. *Nature* 1978, 275 (5679), 416–420, DOI: 10.1038/275416a0.
- Lehn JM. Toward Complex Matter: Supramolecular Chemistry and Self-Organization. *PNAS* 2002, 99 (8), 4763–4768, DOI: 10.1073/pnas.072065599.
- Leibniz GW. *The Monadology and Other Philosophical Writings*. Oxford, The Clarendon Press, 1898.
- Leibniz GW. *La Monadologie*. Paris, Le Livre de Poche, 1991 [1714].
- Lewontin RC. *The Triple Helix: Gene, Organism, and Environment*. Cambridge, Harvard University Press, 2001.
- Liu J, Nussinov R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLoS Computational Biology* 2016, 12 (6), e1004966, DOI: 10.1371/journal.pcbi.1004966.
- Lopez P, Baptiste E. Molecular Phylogeny: Reconstructing the Forest. *Comptes Rendus Biologies* 2009, 332 (2–3), 171–182, DOI: 10.1016/j.crv.2008.07.003.
- Lytard JF. *La Condition postmoderne. Rapport sur le savoir*. Paris, Les Éditions de Minuit, 1979.
- Marks J. Molecular Biology in the Work of Deleuze and Guattari. *Paragraph* 2006, 29 (2), 81–97, DOI:10.1353/prg.2006.0016.
- Maury CP. Origin of Life. Primordial Genetics: Information Transfer in a Pre-RNA World Based on Self-Replicating Beta-Sheet Amyloid Conformers. *Journal of Theoretical Biology* 2015, 382, 292–297, DOI: 10.1016/j.jtbi.2015.07.008.
- Melvin RL, Salsbury FR Jr. Visualizing Ensembles in Structural Biology. *Journal of Molecular Graphics and Modelling* 2016, 67, 44–53, DOI: 10.1016/j.jmgm.2016.05.001.
- Monod J, Changeux JP, Jacob F. Allosteric Proteins and Cellular Control Systems. *Journal of Molecular Biology* 1963, 6, 306–329, DOI:10.1016/s0022-2836(63)80091-1.
- Needleman D, Dogic Z. Active Matter at the Interface between Materials Science and Cell Biology. *Nature Reviews Materials* 2017, 2 (9), 17048, DOI: 10.1038/natrevmats.2017.48.
- Novarina V. *Voie Negative*. Paris, POL, 2017.
- Ovchinnikov S, Park H, Varghese N, et al. Protein Structure Determination Using Metagenome Sequence Data. *Science* 2017, 355 (6322), 294–298, DOI: 10.1126/science.aah4043.
- Palamini M, Canciani A, Forneris F. Identifying and Visualizing Macromolecular Flexibility in Structural Biology. *Frontiers in Molecular Biosciences* 2016, 3, 47, DOI: 10.3389/fmolb.2016.00047.
- Renn JP, Junker M, Besingi RN, Braselmann E, Clark PL. ATP-Independent Control of Autotransporter Virulence Protein Transport via the Folding Properties of the Secreted Protein. *Chem Biol* 2012, 19 (2), 287–296, DOI: 10.1016/j.chembiol.2011.11.009.
- Rout SK, Friedmann MP, Riek R, Greenwald J. A Prebiotic Template-Directed Peptide Synthesis Based on Amyloids. *Nature Communications* 2018, 9 (1), 234, DOI: 10.1038/s41467-017-02742-3.
- Saarikangas J, Barral Y. Protein Aggregation as a Mechanism of Adaptive Cellular Responses. *Current Genetics* 2016, 62 (4), 711–724, DOI: 10.1007/s00294-016-0596-0.
- Sanchez T, Chen DT, DeCamp SJ, Heymann M, Dogic Z. Spontaneous Motion in Hierarchically Assembled Active Matter. *Nature* 2012, 491 (7424), 431–434, DOI: 10.1038/nature11591.
- Scriven M. Explanation and Prediction in Evolutionary Theory. *Science* 1959, 130 (3374), 477–482, DOI: 10.1126/science.130.3374.477.
- Shin Y, Brangwynne CP. Liquid Phase Condensation in Cell Physiology and Disease. *Science* 2017, 357 (6357), eaaf4382, DOI: 10.1126/science.aaf4382.

- Shinde UP, Liu JJ, Inouye M. Protein Memory through Altered Folding Mediated by Intramolecular Chaperones. *Nature* 1997, 389 (6650), 520, DOI: 10.1038/39097.
- Stroberg W, Schnell S. On the Origin of Non-Membrane-Bound Organelles, and their Physiological Function. *Journal of Theoretical Biology* 2017, 434, 42–49, DOI: 10.1016/j.jtbi.2017.04.006.
- Tompa P, Rose GD. The Levinthal Paradox of the Interactome. *Protein Science* 2011, 20 (12), 2074–2079, DOI: 10.1002/pro.747.
- Tompa P, Schad E, Tantos A, Kalmar L. Intrinsically Disordered Proteins: Emerging Interaction Specialists. *Current Opinion in Structural Biology* 2015, 35, 49–59, DOI: 10.1016/j.sbi.2015.08.009.
- Uexküll J von. *Umwelt und Innenwelt der Tiere*. Berlin, Springer, 1909.
- Ullmann A. In Memoriam: Jacques Monod (1910–1976). *Genome Biology and Evolution* 2011, 3, 1025–1033, DOI:10.1093/gbe/evr024.
- Uversky VN. Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *Journal of Biological Chemistry* 2016, 291.13, 6681–6688, DOI: 10.1074/jbc.R115.685859.
- Van Speybroeck L, De Waele D, Van De Vijver G. Theories in Early Embryology. *Annals of the New York Academy of Sciences* 2002, 981 (1), 7–49, DOI: 10.1111/j.1749-6632.2002.tb04910.x.
- Vernadsky V. The Biosphere and the Noosphere. *American Scientist* 1945, 33 (1), 1–12, DOI: 10.1016/j.geoforum.2005.01.004.
- Vickery HB. The Origin of the Word Protein. *YJBM* 1950, 22 (5), 387–393.
- Virchow R. *Archiv für pathologische Anatomie und Physiologie und für klinische Medizin*. Vol. 8. Berlin, Springer, 1855.
- Walsby AE, Hodge MJS. Schrödinger's Code-Script, not a Genetic Cipher but a Code of Development. *Studies in History and Philosophy of Biological and Biomedical Sciences* 2017, 63, 45–54, DOI: 10.1016/j.shpsc.2016.12.002.
- Wasmer C, Lange A, Van Melckebeke H, Siemer AB, Riek R, Meier BH. Amyloid Fibrils of the HET-s (218–289) Prion Form a Beta Solenoid with a Triangular Hydrophobic Core. *Science* 2008, 319 (5869), 1523–1526, DOI: 10.1126/science.1151839.
- Xue B, Romero PR, Noutsou M, et al. Stochastic Machines as a Colocalization Mechanism for Scaffold Protein Function. *FEBS Letters* 2013, 587 (11), 1587–1591, DOI: 10.1016/j.febslet.2013.04.006.
- Zhou M, Guo J, Cha J, et al. Non-Optimal Codon Usage Affects Expression, Structure and Function of Clock Protein FRQ. *Nature* 2013, 495 (7439), 111–115, DOI: 10.1038/nature11833.
- Zimmer C. Scientists are Designing Artisanal Proteins for your Body. *New York Times*, December 26, 2017. (Accessed November 26, 2019, at <https://www.nytimes.com/2017/12/26/science/protein-design-david-baker.html>.)

Sonia Dheur is a molecular biologist, permanent investigator at the CNRS. After conducting research on gene expression and genetic and epigenetic inheritance at IBGC-CNRS UMR 5095, she relocated at UMR5319 Passages (CNRS) where she undertakes critical and reflexive interdisciplinary studies on the recent development of biology as well as studies on common forms of reflexivity in instituted cultural and scientific practices. During the last three years, she has set up and animated a seminar exploring the poetic dimension of scientific production.

Sven Saupe is a geneticist, director of research at the Institut de Biologie et Génétique Cellulaires, CNRS (UMR 5095). In biology, he is interested in genetic and cellular aspects of non-self recognition and cell death processes in fungi. He also undertakes interdisciplinary investigations on certain social, philosophical and symbolic implications of genetics and molecular biology.