When Jacques Monod’s book, Chance and necessity, first appeared in France, it was widely acclaimed for its lucid account of the basic findings and principles of molecular biology. The book also found itself afloat in a sea of political crosscurrents. But neither its value as popular exposition nor its stridently anti-Marxist orientation had little to do with the enormous philosophical ambitions of this essay. Rereading it today, I am filled with admiration for the scope of Monod’s ambition, and for the brilliance of his argumentation. His aim was nothing less than a general theory of living systems, built upon – indeed, made possible by – molecular biology’s theory of the code. It was to finally answer Immanuel Kant’s challenge, to account for the very real peculiarities (or in his term, strangeness) of living beings in a scientific – or objective – manner, which, for him, meant doing so without invoking any notion of final cause. Blind chance, caught on the wing of natural necessity (i.e., of the need to survive) would have to do the trick.

Molecular biology has changed a great deal since those early days, and Monod’s adherence to the dogma so aptly popularized as “DNA makes RNA, RNA makes proteins, and proteins make us” may now seem a bit quaint; certainly
many of the assumptions he made (and defended as unquestionable axioms, as the only possible assumption) appear in hindsight somewhat arbitrary, yet at the same time, the very ingenuity of his efforts to overcome the challenges these early commitments posed can only increase our admiration. Monod could of course not have anticipated the direction future research would take biological thinking about these issues, and I do not want to be understood as suggesting that he might have. Rather, what I do want is to explore the tensions that arose between the various strands of his thought (tensions that are undoubtedly easier to perceive in retrospect), and the ways in which he handled these.

Monod did not doubt that living beings are, without exception, endowed with an aim or project — “in their structure and performance they act projectively, they realize and pursue a purpose” [1]. As Jacob put it, the dream of every cell is to become two cells. To avoid any implication of intentionality or foresight, Monod invoked the term, commonly attributed to Pittendrigh, of teleonomy: end-directedness without final causation. But he well recognized that merely changing the term does not resolve the apparent contradiction; the property of end-directedness remains to be explained.

Teleonomy in fact heads Monod’s list of the defining properties of living beings, the other properties being (b) autonomous morphogenesis, and (c) invariant reproduction. Teleonomy I’ve already explained; autonomous morphogenesis refers to the apparently automatic and spontaneous character of the morphogenetic processes that build the final structure of the organism, i.e. it refers to the part of the sequence that proceeds from proteins to us, or more familiarly in a word, to development. Invariant reproduction refers to the ability of the DNA (the fundamental invariant) to self-replicate and to do so indefinitely, (almost) without error.

Central to Monod’s project was to understand, and to order, these three different properties, and one of his first (and in my mind, more curious) moves was to remove autonomous morphogenesis from his list of essential properties, on the grounds, as he wrote, that “whereas invariance and teleonomy are indeed characteristic ‘properties’ of living beings, spontaneous structuration might rather be considered a mechanism.” (ibid. [1], p. 17).

In other words, once the initial conditions are given, development proceeds as mere mechanics (or physics), and therefore, strictly speaking, it is not in itself biological. Or, to put it differently, it does not contribute to what makes living beings strange. Morphogenesis is a form of self-assembly, like crystal formation — which crucially, he claims, is entirely free of any outside influence. (External conditions may of course impede developmental processes, but they do not govern or guide them). The final structure of a living being is autonomously determined, and morphogenesis simply reveals what was written, further back, in the internal determinism buried deep inside the cell.

What about teleonomy? How are we to understand this property? Enzymes, notes Monod, can function as a kind of Maxwellian demon, exercising an order-creating function. They do this “by virtue of their capacity to form, with other molecules, stereospecific and non-covalent complexes”. Specialized proteins, acting as “detectors and transducers of chemical information, handle the elementary control operations of the cell” (ibid. [1], p. 63), their teleonomic performance resting on their “stereospecific” properties, in turn determined by their amino acid composition. The best example of such apparently cognitive function in proteins is to be found in allosteric enzymes – enzymes in which non-covalent stereospecific binding with other compounds produces a modifying effect on their activity. Both the non-covalence and the stereospecificity of the binding are crucial — the stereospecificity because it allows for recognition, selectivity, and choice; the non-covalence of the binding because of the adjustment between the stability of the complex and the associated function that is made possible by rapid assembly and disassembly. Allosteric enzymes, Monod suggests, function like electronic relays that can integrate the inputs of a number of different sources. In sympathy with his attribution of “cognitive function” to these proteins, I suggest we can think of them as ‘smart molecules’. They are demon-like, first, because of their apparent ability to escape the second law of thermodynamics, and second, because of their functionality, their apparent end-directness.

For Monod, allosteric complexes are microcybernetic adaptive machines, which, like the man-made cybernetic machines we find in the macro world, owe their teleonomic – i.e., end-directed – performance to their makers. Thus, the importance of Monod’s distinction between activity and synthesis: allosteric regulation may be a teleonomic process, but like the thermostat, only because of its particular structure, and this, Monod insists, is due to its amino acid sequence, dictated by the nucleotide sequence of its parent DNA. But here, a small problem arises.

The whole point of allosteric proteins is that the structure is not completely dictated by its amino acid sequence: its quaternary structure depends on its immediate environment, on the presence or absence of other molecules with which it can bind. But at the same time, Monod’s commitment to what I might call the central dogma writ large obliges him to locate the source of the information required for teleonomy, like that required for morphogenesis, exclusively in the DNA where it can be safeguarded and perpetuated without external influence. Thus, e.g., he tells us that the genetic information determines not only the sequence of amino acids but also the conformation of the resulting protein, and it does so exclusively, for the protein assumes its functional shape spontaneously and autonomously. “[A]mong the thousands of different ways in which the polypeptide fiber could theoretically bundle itself”, he writes, “only one is actually adopted. … For the folded form … only a single shape and state actually obtains…With this state and no other its functional activity is connected.” (ibid. [1], p. 92). And later, “[n]o supplementary input of information other than the genetic is necessary; none, it seems, is even possible.” (ibid. [1], p. 109).

Clearly, between Monod’s recognition of the powers of allosteric regulation (he writes, e.g., “so far as regulation through allosteric interaction is concerned, everything is possible” – ibid. [1], p. 77), and his commitment to the
unidirectionality of information flow, a significant glitch has surfaced – a glitch that is frequently papered over by the insertion of two parenthetical insertions to qualify claims that only a single conformation obtains: first, “under normal physiological conditions”, and second, “or at the very most, a small number of discrete states, not very different from each other” (ibid. [1], p. 91). We might of course want to know, if we were so inclined, what counts as a “normal physiological condition” – does it include, e.g., the cellular milieu? And what counts as a small number, or not very different? Why might not these differences be important? But in 1970, few would have been so inclined.

Finally, going down one more level, the question arises, how is teleonomy related invariant reproduction? How is the activity of proteins related to that of DNA? Consistent with the sharp distinction he advocates between synthesis and activity, between the different roles of DNA and protein, Monod now introduces his final hypothesis. He writes:

“Invariance necessarily precedes teleonomy. Or, to be more explicit, the Darwinian idea that the initial appearance, evolution, and steady refinement of ever more intensely teleonomic structures are due to perturbations occurring in a structure which already possesses the property of invariance – hence is capable of ‘preserving the effects of chance and thereby submitting them to the play of natural selection’. . . . Ranking teleonomy as a secondary property deriving from invariance – alone seen as primary – the selective theory is the only one so far proposed that is consistent with the postulate of objectivity. It is at the same time the only one not merely compatible with modern physics but based squarely upon it, without restrictions or additions. In short, the selective theory of evolution assures the epistemological coherence of biology and gives it its place among the sciences of ‘objective nature’ (ibid. [1], pp. 23–24).

A hypothesis, he writes, but it is the only one he regards as possible.

Elsewhere, Monod is sharply critical of those who would restrict the teleonomic principle to living matter. He writes, “[t]he secret of DNA’s invariance resides in the stereo-chemical complementarity of the non-covalent complex constituted by the two strands.” (ibid. [1], p. 105).

Today however it is well recognized that the complementary structure of DNA is not by itself sufficient to guarantee the fidelity of replication, and that the property of invariance as itself an exceedingly complex achievement. Left to itself, DNA would not replicate at all. And even with the necessary polymerases required for replication, without an elaborate system of proofreading, editing and repair, errors would accumulate at such a rate as to undermine any meaning of invariance. A number of different processes are involved in ensuring faithful replication in the course of its occurrence. One works by helping to select the correct nucleotide for complementary binding. Another, by checking the most recently added nucleotide, and immediately removing it if it should fail the test of complementarity. A third comes into action after a new strand has been synthesized, and it works by repair mismatches that might have occurred in spite of the first two “error-avoidance” mechanisms. Mechanisms of “excision repair” – those first observed in the early work on “photo-reactivation” – are of yet a fourth kind. These come into play in response to environmentally inflicted damage (e.g., ultraviolet light), and they work to restore the integrity of the DNA. If the damage has been confined to a single strand, excision repair mechanisms can reverse that damage with little chance of error. Finally, additional mechanisms guaranteeing the global integrity of the chromosome have also been identified. The stability of gene structure thus appears not as a starting point but as an end product – as the result of a highly orchestrated dynamic process requiring the participation of scores of enzymes organized into complex metabolic networks that regulate, and ensure, both the stability of the DNA molecule and its fidelity in replication. As the late Robert Haynes has written, “[t]he stability of genes is now seen to be more a matter of biochemical dynamics, than of the molecular ‘statics’ of DNA structure. The genetic machinery of the cell provides the most striking example known of a highly reliable, dynamic system built from vulnerable and unreliable parts” [2]. In other words, a system which already possesses the property of invariance is, simply by virtue of that fact, already tantamount to a living system, and it arrives on the scene only after eons of evolution.

About evolution, Monod had quite a bit to say. Indeed, it would seem that he regarded evolution, or rather, natural selection, as the only truly biological force. Morphogenesis was physics, and teleonomic performance, microcybernetics (or chemistry), the two together providing the machinery for the revelation of biological form. Only in natural selection, operating on DNA, did he see the possibility for the creation of biological form. Unlike morphogenesis and teleonomy, evolution does not reveal a plan to be found elsewhere; it merely accumulates the radical innovations forged by the operation of natural selection on random changes in the DNA.

“According to modern theory”, he writes, “the idea of ‘revelation’ applies to epigenetic development, but not of course to evolutionary emergence, which, owing precisely to the fact that it arises from the essentially unforeseeable, is the creator of absolute newness.” (ibid. [1], p. 116). Thanks to the replicative structure of DNA, natural selection is able to transduce noise into music, fortuitous perturbations that themselves have no relation whatsoever to teleonomic function (and which in a nonliving system would simply degenerate) into order, function, even purposive beings. “Drawn out of the realm of pure chance, the accident enters into that of necessity.”
(ibid., [1], p. 118). And all this because of the unique self-replicative properties of DNA.

What we have here is of course, at least by now, an abundantly familiar story – it is the standard, neo-Darwinian doctrine as popularized by people like Dawkins and Dennett. What is important, and to me of interest, is Monod’s particular way of integrating the Central Dogma into this doctrine, most especially, the role of the barriers he erects between nucleic acid molecules and proteins, between synthesis and activity and, between reproduction and regulation, creation and revelation, noise and determinism, and finally, between chance and necessity. By his own account, the replicative structure of DNA allows the work of mutation an “unrestricted liberty of creation”. One might think, then, and especially in hindsight, that mutation might generate mechanisms that could feed back onto its own domain, regulating the rate and distribution of changes in nucleotide sequence – ultimately limiting its freedom but in the process of vastly expediting the pace of evolution. But for Monod, such a possibility was excluded a priori, for it would have transgressed the walls of his epistemological structure.

As it happens, virtually all of the dichotomies on which his opus depends have since broken down, beginning with sharp separation between reproduction and regulation. Perhaps it was unrealistic to assume – even as early as 1970 – that base pair complementarity would suffice to insure reproductive invariance. One might even suggest that the in-principle neglect of other forms of inheritance, however compelling the elegance of inheritance by DNA alone, indicated a certain myopia. But surely, no one could have anticipated the discovery of split genes, multiple coding, RNA regulation, etc. Molecular biology has undergone enormous transformations, and the complexities of the regulatory systems that have now been documented would have astounded Monod. Today, the importance of regulation in genetic systems has far outstripped that of coding – indeed, it is estimated that only 1–2 % of the human genome is devoted to coding for amino acid sequences.

What I find so remarkable, and so extremely interesting, is the extent to which, although he could not have foreseen them, Monod actually pioneered many of these developments. He gave us not only our first regulatory system, but also our first smart molecules – i.e., he gave us allosteric proteins. But both of these contributions hung in a certain tension with his primary commitments. Elsewhere I have written about Monod’s efforts to cast the mechanics of regulation in ways that would sustain the notion of “gene action” [3]: here, I would like to focus in the ways in which his ontological commitment to the primacy of DNA constrained his thinking about the power of allostery. Although he wrote that “so far as regulation through allosteric interaction is concerned, everything is possible”, for him, not everything was conceivable. In particular, what was not conceivable was anything that violated his dichotomies.

So, in the remainder of this essay, I want to consider a rather different approach to the problem of the evolution of function, of end-directedness – one that explores the possibilities of one of Monod’s excluded domains, i.e., of evolution before invariant reproduction, and that gives allostery a central role.

In my view, (and I’ve already hinted as much), some sort of teleonomic regulation must have arisen prior to natural selection, for the simple reason that natural selection, at least as conventionally understood, requires the prior existence of stable, autonomous, and self-reproducing entities. Single celled organisms, e.g., or, simply, stable, autonomous, cells capable of dividing. But these first cells were, of necessity, already endowed with numerous subcellular entities (or modules) endowing the primitive cell with the functions minimally required for the cell to sustain itself and reproduce. In other words, even if the first cells lacked many features of the modern cell, they had to have had primitive mechanisms to support metabolism, cell division, etc.; there needed to have come into being primitive embodiments of function that would work to keep the cell going and to protect it from insult.

Such mechanisms came into being not as a result of natural selection but as a consequence of the internal selection that follows automatically from their contribution to the persistence of the system of which they are part. It is a form of selection that does not depend on reproduction (which might be regarded as one way of ensuring persistence, rather like autocatalysis) but rather, a more general kind selection of which natural selection is a particular example. Indeed, their existence is what lends the cell the stability necessary for natural selection to operate.

As Monod himself argued, the existence of such mechanisms is crucial to what makes a system qualify as biological, and the difficult task is to account for how they might have originally come into being. How might such devices – devices that bear all the marks of design – have arisen naturally, without a designer? How can we account for the origin of entities capable of persisting long enough for Darwinian selection to operate? How can we account for the origin of the primordial cell?

Despite all our efforts, the critical properties of function, agency, and purpose continue to mark organisms (even if not machines) apart from the kinds of systems with which physics deals. An account of how properties of this sort might emerge from the dynamics of effectively homogeneous systems of simple elements, however complex the dynamics of their interaction might be, continues to elude us.

Cybernetics, and its emphasis on the relation between feedback and function characteristic of homeostatic devices, offered one clue: I believe that Herbert Simon offered us another. In fact, it is sobering to go back and reread Simon’s 1962 essay on “The architecture of complexity” [4]. Here, Simon introduces a crucial if much neglected argument for a form of evolution that is alternative both to natural selection and to emergent self-organization: evolution by composition. The idea is this: if stable heterogeneous systems, initially quite simple, merge into composite systems that are themselves (mechanically, thermodynamically, chemically) stable, such composite systems in turn can provide the building blocks for further construction. Through repetition, the process gives rise to a hierarchical and modular structure that Simon claims to be
the signature of systems with organized complexity. “Direction”, he explains, “is provided to the scheme by the stability of the complex forms, once these come into existence. But this is nothing more than survival of the fittest – that is, of the stable.” (ibid. [4], p. 191).

We need to be a bit careful here about what we mean by stability – we are not interested in the stability of rocks, and perhaps not even of the limit cycles of dynamical systems closed to informational or material input. Rather, we are interested in the stability of non-equilibrium systems that are by definition open to the outside world, not only thermodynamically but also materially. Perhaps a better word would be robustness. The systems that endure are those that are robust with respect to the kinds of perturbations that are likely to be encountered. The critical questions then become, first, how do new ways of persisting – new stable modes of organization – come about, and second, how are they integrated into existing forms?

In neo-Darwinian theory, novelty arises through chance mutations in the genetic material and is integrated into existing population by selection for the increased relative fitness such mutations might provide. In the picture Simon evokes, novelty arises through composition (or combination), is further elaborated by the new interactions that the proximity of parts brings into play, and, finally, integrated into the changing population by selection for increased relative stability. Of particular importance is the stability of the composite acquired with the passage of sufficient time to undergo a process of mutual co-adaptive changes under the optimizing forces of selection. Symbiosis provides what is probably the best example of all three aspects of the process, and perhaps especially of the ways in which the net effect is to bring into being entirely new kinds of entities that would persist by virtue of their enhanced robustness.

But over the long history preceding the arrival of the first cell, a different kind of composition was required – not composition of existing life forms, but composition of complex molecular structures (like proteins, or nucleic acids, or complexes of these macromolecules). Molecular composition rather than symbiosis. (Or perhaps simply what Jean-Marie Lehn refers to as supra molecular chemistry).

A crucial question is: how do such molecular or supramolecular composites come about? To be sure, random collisions are a big part of the picture. But molecules, and especially large molecules, like proteins, are not billiard balls. They are sticky, they have binding sites, hooks, that actively engage other molecules, that invite the formation of larger complexes through the formation of covalent and non-covalent bonds. There is a kind of inherent activity to such collections of molecules – perhaps bearing some resemblance to what Lamarck sought in his imponderable fluids, but here is activity that has already been internalized – long before the complexities of animal movement could evolve. In many macromolecular complexes, the springs of activity are built into their very structure, they are already internalized. Drawing energy from their interactions with the environment, they are molecular motors. We might speak of chemical forces and free energies rather than of caloric and electricity but the idea seems to me to have distinct echoes in Lamarck’s earlier vision.

In any case, there is another important point about molecular composition that I need to emphasize, perhaps even the most important point. The formation of such covalent and non-covalent bonds can also change the components with which the process started – thereby creating the possibility for new interactions, new binding sites, new hooks. These new capacities are not simply the consequence of the new proximities molecular binding creates, but also of the changes that have been triggered in the component parts. Macromolecules like proteins are not only not billiard balls, but also, they are not simply sticky balls. They are composite structures that are often – perhaps usually – capable of stabilizing in a variety of distinctive shapes, or forms. The binding of other molecules can trigger a shift from one conformation to another, thereby exposing new binding sites, new possibilities for subsequent composition. Prions provide a good example of what I am talking about. Prions are proteins that become infectious agents as a result of a change in folding – a change in folding endows the molecules with the capacity to transmit their new state to other (normally folded) proteins when they come into contact. But this is nothing more than allostery. The important point is, I suggest, that it adds a new dimension to evolution – i.e., to evolution under the pressure for increased stability. Prion infection results from a single allosteric change, but more generally, and especially with the possibility of cumulative changes, allostery provides a mechanism for exploring new evolutionary spaces and for accelerating the formation of ever more complex, and perhaps even functional, structures.

Such processes would seem to be especially pertinent to the evolution of cellularity. Biological cells are replete with devices for ensuring survival, stability, robustness. Think, for example, of the structures (devices) that have arisen to regulate cell division, ensuring that cell division is not triggered too early (when the cell is too small) or does not wait too long (when the cell has gotten too big). Or of the vastly complex kinds of machinery for guaranteeing fidelity in DNA replication, the accuracy of translation, or the proper folding of proteins. Each of these processes – or functions – could presumably have evolved by virtue of the enhanced stability/persistence that the structures on which they depend lend to the system of which they are part.

Because each such mechanism transforms the available options, and pathways, for subsequent evolution, their arrival might be said to mark off different evolutionary epochs. Nucleic acid molecules, for example, appearing on the scene long before the advent of anything like a primitive cell, introduced a significant advance over mechanisms of autocatalysis for making more because it made possible the replication of molecules with arbitrary sequences. The subsequent formation of a translation mechanism between nucleic acid sequences and peptide chains required the combination of already existing nucleic acid molecules and already existing protein structures. But the innovation of a translation
mechanism – in effect, the advent of genes – ushered in an entirely new order of evolutionary dynamics dominated, according to Carl Woese, by horizontal gene transfer. Woese argues that cellular evolution, precisely because it needed so much componentry, “can occur only in a context wherein a variety of other cell designs are simultaneously... [and] globally disseminated”. He writes, “[t]he componentry of primitive cells needs to be cosmopolitan in nature, for only by passing through a number of diverse cellular environments can it be significantly altered and refined. Early cellular organization was necessarily modular and malleable” [5]. Indeed, only with the sealing off these composite structures and the maintenance of their identity through growth and replication – i.e., after a few hundred million years of extremely rapid evolution – did individual lineages become possible. As Freeman Dyson puts it, “one evil day, a cell resembling a primitive bacterium happened to find itself one jump ahead of its neighbours in efficiency. That cell separated itself from the community and refused to share. Its offspring became the first species. With its superior efficiency, it continued to prosper and to evolve separately” [6]. The rest, as they say, is history – i.e., the history of Darwinian evolution.

But long before the advent of that cell, long before anything like the biological cell became possible, another (perhaps almost equally important) transition occurred, and this is the advent of what I might call “smart matter”, or rather, smart molecules. Smart molecules are molecules that can both register (sense) signals in their environment and respond by changing their rules of engagement – e.g., allosteric molecules. I suggest that such molecules came on the scene somewhere over the course of the evolution of macromolecules, like DNA and proteins, and further, that their appearance was crucial to the subsequent evolution of living systems. Ray Kurzweil, undoubtedly employing a somewhat different notion of “smart”, has written that “once matter evolves into smart matter... it can manipulate matter and energy to do whatever it wants” [7]. I would not go quite that far, but I would suggest that once matter evolves into smart matter, the range of what it can do becomes enormously expanded.

Monod banked on random perturbations to generate a world of infinite possibility, but why would evolution have limited itself to such a primitive mechanism? Why restrict the genesis of novelty to the operation of chance? After all, the fruits of evolution’s own operation – either before or after the onset of natural selection – are themselves capable of generating novelty – indeed, of opening up entirely new realms of possibility. The emergence of allosteric molecules is just one example.

References