Multi-scale *vs* function-dependent chromatin modeling

Annick Lesne, Christophe Lavelle & Arndt Benecke

System Epigenomics Group, Institut des Hautes Études Scientifiques & Institut de Recherches Interdisciplinaires – CNRS
35 route de Chartres, F-91440, Bures sur Yvette, France [lesne, lavelle, arndt]@ihes.fr

Abstract

Transcriptional regulation involves structural and dynamic events at many different time and space scales, the latter ranging from local DNA sequence to entire chromosomes. Multi-scale approaches inspired from theoretical physics have been developed in a bottom-up way to integrate parameters and mechanisms at a given scale into effective, and hopefully reduced, descriptions at higher scales. In chromosomes, however, the structures and dynamics at a given scale also exert a strong influence on smaller-scale ingredients, quantitatively controlling and even qualitatively modifying their properties. Self-consistent or iterative 'up-and-down' approaches are to be introduced to account for the strong interconnections between the levels and ensuing circular causal schemes, leading to an endless and prohibitive increase of complexity of the descriptions and models. Adopting another approach, we propose to devise models taking the biological function as a starting point and continued guideline; decomposition is achieved by dissecting its logic and implementation into basic processes. These elementary processes involve features at different scales and are already integrated in their formulation. More generally, such a decomposition results in 'self-scaled' functional modules, independent of the arbitrary description or observation scale. Self-scaling provides the relevant complexity reduction conjunct with the functional descriptive power in modeling complex living systems at multiple scales.

1 From DNA to chromosome: the central role of chromatin

It is now acknowledged that in eukaryotic cells, DNA is rarely present as a bare molecule but rather as a part of a nucleo-protein complex: the *chromatin* [34]. Its basic motif is made of the wrapping of 146 base pairs (bp) around an octamer of histone proteins, regularly spaced all along DNA. This motif is named a *nucleosome* and the DNA stretch connecting two adjacent nucleosomes, of length varying according to the species and cell type (roughly equal to a few tens of bp), is called a *linker*. This first level of genomic organization is observed in all eukaryote species; it is moreover highly conserved, hinting at a key role of nucleosomes in eukaryote mechanisms for implementing gene expression and cell differentiation. Other proteins are involved in chromatin, as well as ions. The assembly is further organized at different scales, up to the entire chromosome (see Figure 1). Among these higher levels of organization a central one, that might be observed in vivo and purified or reconstituted in vitro, is the *chromatin fiber*, also known as the *30 nm fiber* due to its roughly constant diameter of 30 nm. Our investigations on transcriptional regulation (e.g. [5], [19] [13], [6] [17]) and the discussion presented here about the relevant modeling approaches in this context, focus on this chromatin fiber. We argue that it is the central functional articulation between on the one hand, cell metabolism and signaling pathways and on the other hand, gene expression and nuclear regulatory networks [6] [19].

2 Multi-scale view of the chromatin fiber

2.1 A multi-scale organization

A major breakthrough has been provided by both experimental and theoretical investigations focusing on the multi-scale organization of the chromosomes and its mechanistic consequences, in particular for transcription and its control. As illustrated in Figure 1 or similar ones encountered in textbooks [34], structures at different scales superimpose in chromosomes: the DNA molecule, the nucleosomes and associated 10 nm beads-on-string fiber, the 30 nm chromatin fiber, chromatin loops and other (still debated [35]) higher-level structures, up to the entire chromosome.

The simplest multi-scale approach is to dissect the chromosome into *well-identified elementary structures* (DNA, nucleosomes, chromatin fiber) at *well-separated scales* (1 nm, 10 nm and 30 nm respectively) and to thoroughly investigate the properties of these elements in isolation [31]. But these elements at different scales are actually not isolated, and their properties reflect directly (though in an integrated way) in the higher-level features. In the spirit of a famous paper by P.W. Anderson entitled 'More is different' and describing the hierarchy of qualitatively different models (or even theories) that might be involved to investigate one and the same physical object [1], theoretical bottom-up approaches have been developed to derive effective coarse-grained models from more microscopic and more detailed ones. This point is illustrated below in § 2.2 and § 2.3 with studies we have conducted to determine mechanical and topological features of the chromatin fiber from the knowledge of DNA properties.

A first benefit of such a hierarchical bottom-up model is to provide a frame to interpret, exploit and integrate experimental or simulation data obtained at different scales. Let us quote as an illustrative sample some experimental facts and studies that will be discussed further in the present paper: epigenetic information about DNA methylation and

Modelling Complex Biological Systems in the Context of Genomics, pp. 59–70, coordonné par P. Amar, F. Képès, V. Norris, M. Beurton-Aimar et J.P. Mazat, EDP Sciences (2006).

histone-tail post-translational modifications [23], chromatin immuno-precipitation, nucleosome polymorphism [28], DNA [30] and chromatin fiber [12] [2] micro-manipulations, or numerical modeling [21].

But we shall emphasize in § 3 a biological specificity of this multi-scale organization, associated with its evolutionary history. In this regard, 'Life is different' and the analysis of Anderson cannot be straightforwardly transposed from physics to biology [24]. A main difference is the presence, in particular within chromosomes, of specific regulatory schemes, according which hyperstructures at the largest scales influences, controls the lowest scales.

2.2 Structural and kinematic effective models

As mentioned above, we here detail two successful implementations of a bottom-up multi-scale approach, namely the determination of the elastic properties and topological properties of the chromatin fiber, knowing its local architecture and assembly rules, as well as the elastic and topological properties of the underlying DNA stretch.

Elastic properties. Both on experimental (micro-manipulation of a single chromatin fiber [2]) and theoretical grounds [27] [19], the elastic properties of the chromatin fiber can be described within the formalism of *classical mechanics*. Namely, the fiber is seen simply as an *elastic rod* endowed with bend, twist and stretch elastic degrees of freedom. The relation between the stresses (force and torque) applied at the ends of the rod and the ensuing elastic strains in the rod is restricted to its dominant linear contribution, that defines four elastic coefficients [9]: the bend persistence length $\mathcal{A}$, the twist persistence length $\mathcal{C}$, the stretch modulus γ and the twist-stretch coupling factor g. Thermal fluctuations of each molecular degree of freedom (average energy $kT/2$) are accounted for in an effective and integrated way (the elastic coefficients will depend on the temperature T) but the associated stochasticity is no longer described explicitly: at the fiber scale (30 nm), a deterministic model like those developed for macroscopic springs makes sense.

Also DNA can be described as an elastic rod and its elastic coefficients have been both measured experimentally by direct micro-manipulations of a single DNA molecule [30] and computed theoretically [11]. The stretch modulus value larger than 1000 pN shows that the stretch degree of freedom is in fact quenched in physiological conditions (the typical thermal energy kT corresponds to the work performed by a force of roughly 4 pN over a distance of 1 nm), hence it is enough to consider the bend and twist persistence lengths, respectively $A \approx 50$ nm and $C \approx 75$ nm. Note that such a description is in fact averaged over several base pairs and does not explicitly take into account sequence effects, which amounts to consider an homogeneous DNA molecule with no sequence variability.

Similarly, we assume an homogeneous local architecture for the fiber, characterized within a two-angle model [36] by the linker length l and the entry-exit angle ϕ (angle between the two linkers connecting a nucleosome to its neighbors along DNA). It is then possible to determine $\mathcal{A}$, $\mathcal{C}$, γ and g as a function of l, ϕ, A and C. The resulting expression evidences a *strong sensitivity of the fiber elasticity* with respect to these four constitutive parameters, that can be modified both in vivo and in vitro by changes in the ionic concentration, involvement of H1 or non histone proteins, nucleosome repositioning or remodeling, histone-tail post-translational modifications [8] [9]. The chromatin fiber thus appears as a *tunable spring* and its elastic properties cannot be ignored in its in vivo behavior and regulatory functions. Improvements of this minimal model can be considered, accounting for instance for the stacking interactions between the nucleosomes, without modifying the conclusion about the sensitivity of the fiber elastic properties.

Topological properties. The behavior of the fiber as an elastic rod is accompanied with topological properties. These properties are mainly encapsulated in the existence of a topological invariant when considering a fiber stretch with fixed ends (anchored on some substrate, e.g. a MAR region or boundaries, that has the same effect as connecting the ends and forming a closed loop) [20]. This topological invariant L_k^{fiber}, named the *linking number*, corresponds to the number of turns imposed on the fiber (starting from a relaxed state with free ends) before anchoring the ends or closing the loop by gluing these ends one to the other. A remarkable mathematical result relates this quantity to the *twist* T_w^{fiber} of the fiber (number of turns onto itself, around its own axis) and its *writhe* W_r^{fiber} (a quantity related to its path in the three-dimensional space): $L_k^{fiber} = T_w^{fiber} + W_r^{fiber}$. The balance between T_w^{fiber} and W_r^{fiber} varies with the three-dimensional conformation of the fiber, passing from $W_r^{fiber} = 0$ and $L_k^{fiber} = T_w^{fiber}$ in a straight fiber to a torsionally relaxed plectonemic conformation with $T_w^{fiber} = 0$ and $L_k^{fiber} = W_r^{fiber}$.

Such an inter-conversion presumably occurs during condensation and decondensation: it seems indeed possible to trigger decondensation at $L_k^{fiber} = \text{const.}$ by modifying the twist of the fiber through a change in the nucleosome shape [21]. Such a theoretical scenario is made quantitative by the analysis of the topological changes experienced by the underlying DNA molecule. It is indeed to be emphasized that topological properties and invariants can be defined both at the DNA and fiber levels, within the parallel elastic-rod modelings of DNA and the fiber encountered in describing their elastic properties. In particular, the same object is associated with two different notions of linking number: L_k^{fiber} and L_k^{DNA}, according to the considered scale, and one can show that $L_k^{fiber} = L_k^{DNA} + \text{const.}$ [3].

These two examples illustrate how the same issue for the same object can be tackled within different models according to the scale of description (also termed the *averaging scale* since it puts a limit between the details explicitly described and those, of smaller scale, accounted for in an effective average fashion). Considering one rather than the other model depends on the response properties to be analyzed and the nature of the stimuli exerted on the objects, i.e. whether the external stresses and boundary conditions are applied at the DNA or fiber level.

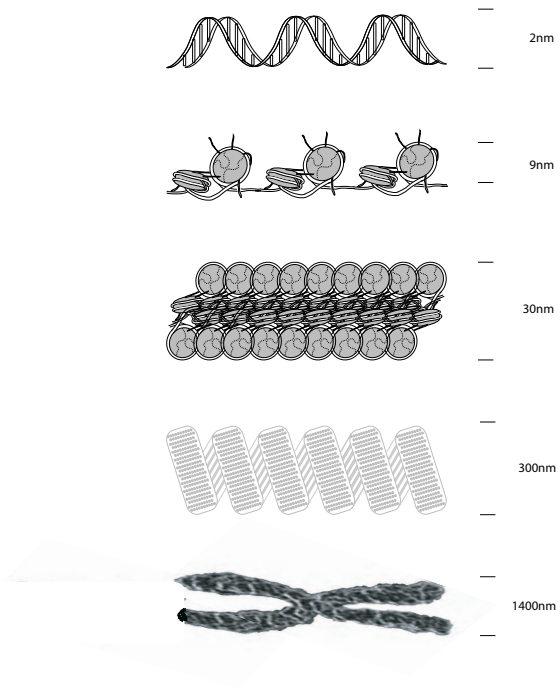

Figure 1: Hierarchical organization from DNA to chromosome, identifying well-defined structures at separated scales: DNA double-helix, nucleosomes and beads-on-string fiber, 30 nm chromatin fiber, chromatin loops, up to the entire chromosome

2.3 Chromatin dynamics and stochasticity

In the previous sub-section, we have described two examples of a non trivial multi-scale modeling, focusing on the connections between structural or kinematic models at different scales (respectively DNA and chromatin fiber scales) and working out the relation between their effective parameters. Similar bottom-up connections can be derived for chromatin dynamics. The cross-level couplings result in remarkable emergent properties (i.e. absent when the elements where they occur are considered in isolation). For instance, coordinated enzymatic histone-tail post-translational modifications generate alternating conformational changes at the chromatin fiber level [4] [5]. The associated '*chromatin breathing*' involves several nucleosomes and exhibits a far supra-molecular period of a few tens of minutes, evidencing the increase in the spatial and temporal scales of relevance for observing this emergent phenomenon, although the basic events occur at molecular scale.

Another bottom-up issue is to determine and describe the *level of stochasticity* present at the different scales and its functional consequences. The basic stochasticity originates in thermal fluctuations. In case of independent fluctuations, the resulting behavior at higher scales is dominantly deterministic, since accumulating independent fluctuations average out according to the law of large numbers. Observing or invoking stochastic events at large scales thus requires strong coupling between the molecular mechanisms, or strong correlations between the successive events, that generate an anomalous large-scale behavior with non negligible fluctuations [15].

2.4 A complex regulatory scheme

Multi-scale chromosome organization does not lead only to bottom-up relationships: the very complexity of living systems and biological functions, in particular the chromosomes and transcription, lies in the presence of *feedbacks from upper scales onto elements at smaller scales that have settled in the course of evolution.* As an illustration, let us cite the control exerted by the chromatin fiber superstructure onto the protein-DNA binding events in linker DNA. Indeed, when a linker is embedded in a condensed chromatin fiber, with stacked nucleosomes acting as fixed anchoring points, its linking number L_k^{linker} is conserved. This topological constraint has energetic consequences, since any strain (e.g. twist or bend) experienced by the linker, in particular the strains generated upon protein binding (e.g. intercalation [33]) should be compensated so as to preserve $L_k^{linker} = \text{const.}$, at some elastic energy cost. The net effect is a noticeable modification of the protein-binding energy landscape. This landscape modification is moreover controlled by any means for tuning the nature of linker anchoring onto the nucleosomes and ensuing tolerance in the linking number constancy; possible means are changes in the chromatin fiber conformation and compaction state, histone-tail post-translational modifications, presence of ions or linker histones.

The strong and evolutionary adapted influence of larger scales onto elements at smaller scales within a chromosome led us to propose a notion of *generalized allostery* [33] [19], referring with this term to conformational transitions and associated change in activity and function, induced by an effector acting at a remote site and with no direct physico-chemical link with the activity. We here suggest that the involvement of this specific effector does not follow from an inescapable physico-chemical law but from a mutual adaptation settled in the course of evolution. In support and illustration of this generalized notion, DNA and chromatin can exhibit an allosteric behavior insofar as

the hyperstructure in which they are embedded (respectively the chromatin fiber or a chromatin loop) and topological frustration it generates, can induce a bistable behavior in their conformation, controlled through modifications of the hyperstructure and the mechanical constraints it induces at lower levels. For instance, linker DNA can pass from a straight conformation to a buckled one with different binding affinity towards intercalating proteins [33]: the buckled conformation appears as the 'active' form of the allosteric linker; the transition to this active form is triggered by the first DNA-binding event and controlled through the tuning of DNA anchoring onto the nucleosomes. Our generalized notion includes as a special case the notion of *nested allostery* in which a cascade of allosteric behaviors in nested sub-systems is initiated by the effector-binding event [26]. It also embeds extensions of the original notion of allostery, e.g. the localized allostery observed in large assemblies and involving only a sub-system delineated by mechanical constraints [32].

This discussion in fact faces a very general property: the embedding in a superstructure can modify the very individual potentialities of low-level elements, hence *precluding a plain bottom-up strategy*. In a similar spirit, turning to more general interactions and couplings than mechanical or structural ones, it is a whole gene network that underlies transcriptional regulation, moreover exerting feedbacks on its own nodes (a node represents here at the same time a gene and the protein it codes): presumably, some nodes will have different individual properties when embedded in the whole network. For instance, interactions experienced by a protein within the gene network might modify its chemical or physical properties, and allow it to establish different interactions than those observed in vitro, in isolation.

3 A necessary shift of paradigm

3.1 Limits of a multi-scale description

We have seen in the previous section that in order to properly account for the autonomous regulation and behavior of a biological system, a multi-scale approach should tackle jointly all the scales, with no way to a priori ignore some microscopic detail [19]. Obviously, any such proper modeling would rapidly reach high level of complexity, the higher the more faithful and realistic the model is, hence *ultimately intractable* [13] [14]. To circumvent this difficulty, we claim that both integrated modeling and supervised data analysis *should parallel the biological functional logic*.

The hierarchy of spatial scales should enter the scene only because these nested scales correspond to our different but all subjective views on the system and to our various experimental accesses, confined in current practice to a given scale. Data analyses and model predictions have to be ultimately bridged with this hierarchical categorization and the integrated model itself should embed, as special restricted situations, models developed at well-defined scales.

More conceptual objections also arise. In particular, scale-wise description is currently associated with a step-wise, sequential conception of the processes and their time course, e.g. chromatin fiber assembly [25] or transcription initiation. Traveling the hierarchy of scales is even logically correlated with temporal progression along a sequence of events where each event is causally rooted in the previous one. The above conclusions on biological complexity and associated causal loops suggest that this view is presumably too naive for biological functions. In the case of chromatin fiber assembly, it properly follows our ways of reconstituting a chromatin fiber but misses essential adaptive dynamics and self-organization occurring in vivo to achieve in one and the same step various delicate balances between the energetic, steric and topological constraints (reflecting e.g. in the nucleosome positioning and individual conformation), to control multi-stable pathways, or to manage with competing chemical reactions or concerted factor recruitment. The sequential view on chromatin dynamics have to be replaced with an *interaction network viewpoint*. It appears necessary to account of the joint dynamic behavior of numerous interconnected elements, for instance DNA sites and their chromatin surrounding, nucleosomes, or cofactors and coregulators [6] [18]. Hence, not only the spatial and structural models but also the dynamic descriptions should evolve into more integrated models.

Note finally that in biology, by contrast to physics, there is no ubiquitous relation between time and space scales, namely the events occurring at the smallest spatial scales are not necessarily the fastest ones (think for instance to ionic currents across membranes or molecular motor motion). This remark further evidences that a *splitting according the scales has no intrinsic relevance* for biological systems.

3.2 Need of a self-scaled, interconnected, function-dependent modeling

The above-discussed limitations and caveats about multi-scale descriptions lead us to propose a drastic change in the paradigm underlying chromatin modeling and more generally the modeling of biological systems, namely a shift from a scale-dependent focus to a *function-dependent* focus. A function-based analysis is required to account, through the notion of function, of the embedding of living systems or sub-systems *within an evolutionary history*. Speaking of the 'functions' of a biological system (by contrast to the 'properties' of a physical system) is a short-cut underlining the *cross-level consistency of all involved mechanisms, following from their co-evolution and resulting mutual adaptation.* Here, living systems meet the artificial ones, with the striking difference that living systems, following millions of years of trials and selective improvements, achieve in general an extraordinary efficiency. They possibly involve a different logic than the one at work in man-made functional and regulatory schemes.

Another motivation to change drastically the descriptive and modeling framework is related to its required *robustness*. Indeed, any modeling (and in fact any description) involves subjective restrictions, approximations, reference points ... and the only way to justify this arbitrary part of the modeling is to show the robustness of the predictions, namely that they do not depend on these subjective choices. In such a prospect, a model based on a multi-scale decomposition is not robust (except in some specific scale-invariant instances rarely encountered in biology): it explicitly involves the chosen scales, the details at each scales, the mechanisms considered at a given scale, and finally the coarse-graining procedures and closure approximations used in relating one scale to the other. On the contrary, a function-based viewpoint, relying on an objective biological fact, will be robust, at least qualitatively, with respect to the modeling choices. Quantitative agreement between predictions and observations might be then obtained with no change in the functioning principles nor in the logical/causal scheme, only requiring a tuning of the parameter values.

Figure 2: Within a function-based approach, all scales are to be considered jointly and the focus shifts to their functional interconnections: each level influences and is influenced by both lower and higher levels, according to causal loops settled in the course of evolution. This hallmark of complexity evidences the need of alternative approaches, replacing multi-scale models with a function-dependent, self-scaled decomposition in elementary processes.

This novel modeling approach, beyond being function-based, is also termed to be self-scaled and interconnected. By '*self-scaled*', we mean that the physical and experimental hierarchy of scales (Å atomic scale, nm base-pair scale, 10 nm bead-on-string, 30 nm chromatin fiber, 300 nm chromonema, up to the micron scale of the chromosomes and entire nucleus) is not an intrinsic and functional feature of the transcriptional regulation and more generally of any nuclear function. In consequence, it it does not provide a relevant nor operational decomposition. Rather, any biological function, in particular the chromatin function and still more specifically transcriptional regulation, has its own way to embed in real space and unfold its elementary processes *across these scales*, achieving *information circulation* in order to perform the processing of genetic and chromatin code and the concrete implementation of regulatory mechanisms, *each at the most convenient scale* (lowest energy cost, availability of co-regulators, adequate uptake mechanisms ensuring the stationarity of the fluxes ...). The biological function also demands an integrated implementation, what we summarize in the term '*interconnected*' (levels). In short, *functions and associated information processing pervade and exploit all scales jointly in their own evolutionary adapted way*, and the design of our modeling should account for and take the largest benefit of this point (Figure 2).

3.3 Basic elementary processes in transcription

Complexity reduction in living systems studies is presumed to rely on the identification of modules. Obviously, such a program makes sense *only if independent or weakly coupled modules can be delineated*. The precise requirement is that the behavior associated with a module is robust, qualitatively independent of the surroundings and inputs; it possibly affects quantitatively the model predictions but not the logical/causal scheme, kind of behavior and regulatory mechanisms. We have demonstrated in § 2 that defining modules as being well-defined structures each observed at a given scale does not fulfill this requirement. We thus propose another way of dissecting the biological system under consideration (here the chromosome) neither according to the experimentally isolable entities (nucleotides, nucleosomes, ...) nor according to the space and time scales, but according to the function [14].

The main idea is to consider basic processes articulating several physico-chemical mechanisms and unfolding across the scales, in a function-oriented scheme involving effective inputs and outputs, so as to get elementary 'responsive' or 'active' building blocks [14]. We term these multi-level and operationally dedicated mechanisms '*basic elementary processes*'. They intend to provide elementary links of an *effective network* achieving the same function and the same regulatory control. They are 'elementary' in the sense of being indecomposable, i.e. meaningful and describable only as a whole. We again underline that they fundamentally differ, in their spirit, of modules introduced as building blocks of well-defined scale. Their description, modeling or experimental observation involve several scales and might extend in time, hence essentially crossing the standard levels of description and thinking. They describe elementary steps of information processing, transformation, circulation. Actually, we might consider a hierarchy by grouping the most elementary processes of our library into more complex ones, still non-autonomous (hence being only a part of a function) up to completion into a function (that can be defined formally, if necessary, as a pattern of

information processing, involving mechanisms and parameters at different scales). A function is adaptive and robust, in the sense of being context-independent: it should be properly achieved in various external conditions and internal states. On the contrary, elementary processes will be context-dependent so as to maintain the function they participate in. Elementary processes (and even more the values of their parameters) are precisely highly sensitive to any external stress, changes in the surroundings or in the cell internal state, in order that the function is robust.

Let us implement more explicitly this program in the context of transcription. The base-pair, DNA and chromatin-fiber well-separated levels are replaced by the following processes:

— *DNA bending*, including its consequences on DNA affinity for various proteins [22];

— *chromatin tethering*, including as a consequence the invariance of the linking number L_k^{fiber} of the end-tethered chromatin loop, and also the invariance of the linking number L_k^{DNA} of the underlying DNA stretch; it follows that any imposed strain generates mechanical constraints in DNA and modifies its protein-binding energy landscape [33];

— *histone-tail post-translational modifications* with both chemical repercussions (in terms of recognition and recruitment of specific factors [29] [6]) and physical repercussions, either electrostatic (change of the local charge density) or mechanical (change of the local anchoring of DNA onto the nucleosome hosting modified histone(s) [33]);

— *hypercycles* involving the coordinated alternation of two reverse enzymatic reactions, e.g. acetylation and deacetylation; they prescribe a rhythm, i.e. a kind of internal clock inside the chromatin fiber, and provide the very first, non specific step of transcriptional initiation, termed chromatin breathing [4] [5].

— *DNA allostery*, where the chromatin fiber hyper-structure induces bistability properties at the DNA level (or more generally a switch potentiality following from the frustration induced by mechanical or topological constraints) controlled by the fiber compaction and triggered by DNA binding events or transactions [19] [17].

4 Tackling complexity in epigenomics

4.1 Several explanatory schemes

For any biological structure or process, several explanations and several levels of causality actually coexist, demanding to be articulated. Indeed, biological systems can be explained at the same time by a set of correlated mechanistic steps or by invoking an ecological (or even 'economical') balance of inputs and outputs, within an adaptive, evolutionary perspective. To take a simile, switching on a car into motion involves both the specific contact key and general thermodynamic principles of motor functioning. The relevant explanation will thus depend on the perspective, either evolutionary, mechanistic or therapeutic; in other words according to the causal representation to be provided, either in terms of the whole evolutionary consistency of the organism, of the outputs at a given point or response to a prescribed perturbation.

For instance, an allosteric reaction can be described at the molecular scale, as a succession (possibly highly complex, e.g. networked) of molecular interactions and modifications, following physico-chemical laws. It might also be seen as a trick invented, stabilized and improved in the course of evolution. Our approach precisely *aims at reconciling mechanistic and evolutionary explanatory frame.* Its function-dependent focus refers to the inescapable natural selection, while the investigations of processes brings back to physic-chemical mechanisms, dissipative structures and self-organization. Additional schemes, for instance the information-theoretic notions of *genetic code* and *genetic program* have also to be considered to capture a comprehensive understanding of, say, cell differentiation in all its complexity [14]. As a way to bridge these complementary explanatory within a novel frame, we are developing a notion of *chromatin code*, reminiscent of the corpus of ciphers, codes, grammars, languages, computations and calculus developed in (theoretical, abstract) computer science, but with an architecture and logic of its own [6] [17].

4.2 Inverse renormalization-group

The strong interplay between the scales described in the previous section, where possibly some microscopic details directly influencing, or influenced by, the structure and dynamics at large scale, is reminiscent of situations encountered in physics, known as *critical phenomena*, where all scales are relevant and do not separate. In this context, a plain bottom-up representation of causality is questionable since the couplings and assembly of processes are strongly nonlinear and intricate: the assembly of two objects each described by a quantity X is not described by $2X$ but by $a(X)X$. A way out such difficulty is a *renormalization procedure* in which all the direct and indirect couplings and correlations are accounted for as a whole, through an effective, integrated and self-consistent contribution, here $a(X)X$ with $a(X) \neq 2$. More generally, renormalization-group offers a way to qualitatively understand and quantitatively predict anomalous behaviors associated with criticality [15]. For biological systems, criticality is even stronger and far more difficult to tackle since macroscopic features might exert some feedback on the very properties and behavior of the smallest scale elements, hence requiring some 'top-down' analysis. Plain top-down representation is confronted to

an issue of information lack, since we try there to infer a detailed description from coarser ones with (far) less degrees of freedom.

To solve these difficulties, we are developing a novel mathematical framework, that we termed '*inverse renormalization-group*, intending in particular to provide a mathematical zoom to describe and understand biological functions and regulatory schemes [7]. Its design is parallel to the biological logic and implementation, insofar as it relies on the main theoretical principle of biology, namely evolution, natural selection and ensuing optimized fitness of biological systems. The information missing in top-down approaches is to be injected through a biological, evolutionary, adaptive, consistency argument, assessing that the microscopic level is so consistently related (co-evolved, mutually adapted and optimized) to the macroscopic level that the very existence of the macroscopic level brings information on the underlying ones. Thus, in adaptive systems (either living or man-made) evolution and selection allow a top-down inference that is essentially different from the inference scheme relevant in physics. Tackling complexity is precisely achieved here, in the reduction by ways of some optimization argument (here inter-level consistency and adaptive robustness) and educated microscopic implementation, of the whole system.

Additional features of biological systems motivate the development of this novel mathematical formalism [7]. We have yet underlined that in the set of biological functions or in the set of elementary processes from which they are built arise sub-categories, associated with various levels of emergence (for instance raw *vs* effective parameters, single-scale *vs* emergent elementary processes ...). A major conclusion follows: there is *no way out a continuous view* with fuzzy boundaries between the levels, the elements, and the elementary mechanisms or events. Moreover, biological systems do not exhibit a 'natural' underlying real space-time structure, but rather *several superimposed proper times* (that of cycles and hypercycles; that of evolution) and *several superimposed intrinsic topologies* (that of development –curved; that of regulatory networks – infinite-dimensional; that of constraints – tensegrity) with competition (hence possibly frustration if beneficial, e.g. to generate multi-stability and switches) or cooperation (hence synergy).

5 Conclusion

Given a biological system (or even yet a physical or chemical one), the proper model is devised to answer a well-posed question. It cannot intend (and should not) to provide a universal representation of the system, nor to fully capture all its properties whatever their nature, scale and context. Otherwise it would resemble the scale-1 empire map described by Borges [10]. Modeling hence requires to focus on some elements and facts while ignoring a large amount of information of no direct relevance for the addressed issue. The merit of a good and efficient model is precisely to extract and enlighten those, and only those features of the system that are presumably relevant for solving the limited, well-delineated issue under consideration. Multi-scale viewpoint is at the same time required to design experiments or to analyze and interpret data, but irrelevant due to the essential cross-scale ('up-and-down') coupling following from evolution. Inspired by approaches currently followed in theoretical physics, this sorting of relevant information is usually done in the real time and space, according to the scales. We have proposed here to rather base this a priori sorting according to the biological functions. This leads to a drastic change in the paradigm underlying the design of experiments and choice of model systems, supervised data analyses and modeling, namely *a shift from a scale-dependent focus to a function-dependent focus*.

References

[1] Anderson, P.W. (1972) More is different: broken symmetry and the nature of the hierarchical structure of science, *Science* **177**, 393–396.

[2] Bancaud, A., Conde e Silva, N., Barbi, M., Wagner, G., Allemand, J.F., Mozziconacci, J., Lavelle, C., Croquette, V., Victor, J.M., Prunell, A., and Viovy, J.L. (2006) Structural plasticity of single chromatin fiber revealed by torsional manipulation, to appear in *Nature Struct. Mol. Biol.*

[3] Barbi, M., Mozziconacci, J., and Victor, J.M. (2005) How does the chromatin fiber deals with topological constraints, *Phys. Rev. E* **71**, 031910.

[4] Benecke, A (2003) Chromatin dynamics are a hyperstructure of nuclear organization, in *Proceedings Modelling and Simulation of Biological Processes in the Context of Genomics*, Amar P., Kepes F., Norris V., and Tracqui, P. (eds.) Platypus Press, pp. 31–40.

[5] Benecke, A. (2003) Genomic plasticity and information processing by transcriptional coregulators, *ComPlexUs* **1**, 65–76 — see also published comment: Buchanan, M. (2003) *ComPlexUs* **1**, 65–69.

[6] Benecke, A. (2006) Chromatin code, local non-equilibrium dynamics, and the emergence of transcription regulatory programs, to appear in *Eur. Phys. J. E*, Focus Point on "Structure and Dynamics of DNA".

[7] Benecke, A. and Lesne, A. (2006) Inverse renormalization-group, in preparation.

[8] Ben Haim, E., Lesne, A., and Victor, J.M. (2001) Chromatin: a tunable spring at work in our chromosomes, *Phys. Rev. E* **64**, 051921.

[9] Ben Haim, E., Lesne, A., and Victor, J.M. (2002) Adaptive elastic properties of chromatin fiber, *Physica A* **314**, 592–599.

[10] Borges, J.L. (1972) On the impossibility of drawing a map of the empire on a scale of 1, in *A universal history of infamy* [Historia universal de la infamia] (1935).

[11] Bouchiat, C. and Mezard, M. (1998) Elasticity model of a supercoiled DNA molecule, *Phys. Rev. Lett.* **80**, 1556–1559.

[12] Cui, Y. and Bustamante, C. (2000) Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure, *Proc. Natl. Acad. Sci. USA* **97**, 127–132.

[13] Lavelle, C., Benecke, A. (2006) Chromatin physics: replacing multiple, representation-centered descriptions at discrete scales by a continuous, function-dependent self-scaled model, to appear in *Eur. Phys. J. E*, Focus Point on "Structure and Dynamics of DNA".

[14] Lavelle, C. and Benecke, A., and Lesne, A. (2006) Function-dependent self-scaled interconnected chromatin models: integrated functions as a combination of basic elementary processes, in preparation.

[15] Lesne, A. (1998) *Renormalization methods*, Wiley, New York.

[16] Lesne, A. (2006) Multi-scale approaches, article MS101 in *Encyclopedia of Mathematical Physics*, J.P. Françoise, G. Naber, and T.S. Tsun (eds.) Elsevier.

[17] Lesne, A. (2006) Chromatin code: beyond an histone code, to appear in *Eur. Phys. J. E*, Focus Point on "Structure and Dynamics of DNA".

[18] Lesne, A. (2006) Complex networks: from graph theory to biology, submitted to *Letters in Mathematical Physics*.

[19] Lesne, A. and Victor, J.M. (2006) Chromatin fiber functional organization: some plausible models to appear in *Eur. Phys. J. E*, Focus Point on "Structure and Dynamics of DNA",
— see also the commentary of this paper by H. Schiessel in the same issue.

[20] Mozziconacci, J., Lavelle, C., Barbi, M., Lesne, A., and Victor, J.M. (2006) A physical model for the condensation and decondensation of eukaryotic chromosomes, *FEBS Letters* **580**, 368–372.

[21] Mozziconacci, J. and Victor, J.M. (2003) Nucleosome gaping supports a functional structure for the 30 nm chromatin fiber, *J. Struct. Biol.* **143**, 72–76.

[22] Parvin, J.D., McCormick, R.J., Sharp, P.A., Fisher D.E. (1995) Pre-bending of a promoter sequence enchances affinity for the TATA-binding factor, *Nature* **373**, 724–727.

[23] Peterson, C.L. and Laniel, M.A. (2004) Histones and histone modifications *Curr. Biol.* **14** R546–R551.

[24] Polanyi, M. (1968) Life's irreducible structure, *Science* **160**, 1308–1312.

[25] Ridgway, P. and Almouzni, G. (2001) *J. Cell Sci.* **114**, 2711–2712.

[26] Robert, C.H., Decker, H., Richey, B., Gill, S.J., and Wyman, J. (1987) Nesting: Hierarchies of allosteric interactions, *Proc. Natl. Acad. Sci. USA* **84**, 1891–1895.

[27] Schiessel, H. (2003) The physics of chromatin, *Journal of Physics - Condensed Matter* **15**, R699–R774.

[28] Sivolob, A., Lavelle, C., and Prunell, A. (2003) Sequence-dependent nucleosome structural and dynamic polymorphism. Potential involvement of histone H2B N-terminal tail proximal domain, *J. Mol. Biol.* **326**, 49–63.

[29] Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications, *Nature* **403**, 41–45.

[30] Strick, T.R., Allemand, J.F., Bensimon, D., Bensimon, A., and Croquette, V. (1996) The elasticity of a single supercoiled DNA molecule, *Science* **271**, 1835–1837

[31] van Driel, R., Fransz, P.F., and Verschure, P.J. (2003) The eukaryotic genome: a system regulated at different hierarchical levels, *J. Cell Science* **116**, 4067–4075.

[32] Van Holde, K.E., Miller, K.I., and Van Olden, E. (2000) Allostery in very large molecular assemblies, *Biophysical Chemistry* **86**, 165–172.

[33] Victor, J.M., Ben Haim, E., and Lesne, A. (2002) Intercalation and buckling instability of DNA linker within locked chromatin fiber, *Phys. Rev. E* **66**, 060901.

[34] Wolffe, A.P. (1998) *Chromatin: structure and function*, Academic Press, New York.

[35] Woodcock, C.L. and Dimitrov, S. (2001) Higher-order structure of chromatin and chromosomes, *Curr. Opin. Genet. Dev.* **11**, 130-135.

[36] Woodcock, C.L., Grigoryev, S.A., Horowitz, R.A., and Whitaker, N.A. (1993) Chromatin folding model that incorporates linker variability generates fibers resembling the native structures, *Proc. Nat. Acad. Sci. USA* **90**, 9021–9025.