Chromatin fiber allostery and the epigenetic code ‡

Annick Lesne[1,2,3], Nicolas Foray[3,4], Guy Cathala[2,3] Thierry Forné[2,3], Hua Wong[3,5] and Jean-Marc Victor[1,2,3,*]

[1] Laboratoire de Physique Théorique de la Matière Condensée, CNRS UMR 7600, UPMC Univ. Paris 06, Sorbonne Universités, F-75005, Paris, France.
[2] Institut de Génétique Moléculaire de Montpellier, CNRS UMR 5535, Univ. Montpellier, F-34293, Montpellier, France.
[3] CNRS GDR 3536, UPMC Univ. Paris 06, F-75005, Paris, France.
[4] INSERM UMR1052 — Radiobiology Group, Cancer Research Centre of Lyon, F-69008, Lyon, France.
[5] Institut Pasteur, F-75015, Paris, France
* Corresponding author. LPTMC case courrier 121, Université Pierre et Marie Curie, 4 place Jussieu, F-75252 Paris cedex 05, France. Tel: +33 144272666. Fax: +33 144275100. Email: victor@lptmc.jussieu.fr

Abstract. The notion of allostery introduced for proteins about fifty years ago has been extended since then to DNA allostery, where a locally triggered DNA structural transition remotely controls other DNA-binding events. We further extend this notion and propose that chromatin fiber allosteric transitions, induced by histone-tail covalent modifications, may play a key role in transcriptional regulation. We present an integrated scenario articulating allosteric mechanisms at different scales: allosteric transitions of the condensed chromatin fiber induced by histone-tail acetylation modify the mechanical constraints experienced by the embedded DNA, thus possibly controlling DNA-binding of allosteric transcription factors or further allosteric mechanisms at the linker DNA level. At a higher scale, different epigenetic constraints delineate different statistically dominant subsets of accessible chromatin fiber conformations, which each favors the assembly of dedicated regulatory complexes, as detailed on the emblematic example of the mouse Igf2-H19 gene locus and its parental imprinting. This physical view offers a mechanistic and spatially structured explanation of the observed correlation between transcriptional activity and histone modifications. The evolutionary origin of allosteric control supports to speak of an "epigenetic code", by which events involved in transcriptional regulation are encoded in histone modifications in a context-dependent way.

PACS numbers: 87.15.-v, 87.16.Sr, 87.15.La, 81.16.Fg

Keywords: epigenetics, chromatin fiber, histone-tail acetylation, allosteric transition, transcriptional regulation

Abbreviations: DNA, deoxyribonucleic acid; H3, the histone type H3; H4, the histone type H4; H4K16, a specific residue (lysine 16) of histone H4; HAT, histone acetyltransferase; HMG, high-mobility group (a group of proteins); TF, transcription factor; tRNA, transfer ribonucleic acid.

‡ This work has been funded by the French Institut de la Recherche Médicale, grant MICROMEGAS

1. Introduction

The notion of allosteric interaction has been introduced for proteins in 1961 [42]. Since then, countless works have demonstrated its importance, see for instance the recent overview [21]. The main features of protein allostery are:
– action at a distance, where a signaling event at an effector site triggers or facilitates a functional event at a different site of the protein, currently termed the active site when the protein is an enzyme. Allostery thus achieves an increase of the interaction range far beyond atomic scale.
– a reversible conformational switch of the allosteric entity triggered by the effector binding and modifying the active site; this switch as a whole, linking the effector and the active sites, is ensured by the cohesive structure and stiffness of the allosteric protein;
– gratuity, insofar as allostery allows regulation of any enzyme by any molecule potentially binding the enzyme and affecting its conformation with a large enough effect [41]. In this respect, the effector acts as a signal having no necessary chemical relation to the enzyme catalytic activity [40]. Accordingly allostery plays a central role in signal transduction [48].

Two instances are encountered: homotropic allostery corresponds to the case where the effector site and the active site are structurally identical, and bind molecules of the same species; cohesiveness then typically originates from the symmetry of the protein subunits, as in the emblematic example of haemoglobin [50, 51]. Heterotropic allostery refers to a more general situation where the effector and the ligand are different molecules binding a possibly monomeric protein. In short, homotropic allostery is associated with cooperative binding, whereas heterotropic allostery is associated with signaling.

Allostery and cooperativity should be clearly considered as different concepts: cooperativity does not necessarily originate in an allosteric mechanism, and allostery does not necessarily reflect in cooperativity. Cooperative kinetics, namely a Hill kinetics replacing the plain Michaelis-Menten kinetics, have been associated with homotropic allostery in the seminal model by Monod, Wyman and Changeux, which has been developed to explain the cooperativity of ligand binding in enzymatic catalysis, in this case oxygen binding to haemoglobin [43]. Cooperativity here means that the cost per binding event decreases with the number of binding sites, as a result of a concerted transition to an active form of the subunits hosting the sites. But the notion of allostery is far more general, including instances of heterotropic allostery that are not related to cooperativity and do not display a specific kinetic signature. Conversely, cooperativity is not necessarily due to an allosteric mechanism, and may for instance originate in electrostatic collective effects (see Section 2.4 below). We here favor a mechanistic view, at the single molecule level, in terms of structural or conformational transitions rather than in terms of concentrations and kinetic rates. Allostery, in its basic and original meaning, is the action at a distant active site of a structural change in the allosteric entity induced by the binding of an effector molecule. In this perspective, allostery is more reliably associated with facilitation and coordination. Namely, the conformational transition allosterically triggered by the effector event facilitates further events, typically by a change in ligand binding affinity

PC201104, by the French Institut National du Cancer, grant INCa_5960, and by the French Agence Nationale de la Recherche, grant ANR-13-BSV5-0010-03. UPMC belongs to Sorbonne Universités.

that can be assessed on thermodynamics grounds [11] or at the atomic level [18]. It thus mechanically coordinates the effector and active sites.

Surprisingly, the scope of allostery has been barely extended to entities other than proteins. Only a few papers [3, 10, 12, 55] mention the possible allosteric behavior of DNA. We here go further and consider the functional role of allosteric transitions of the chromatin fiber. There are several experimental evidences that transcriptional regulation in the eukaryotic realm is not only a matter of transcription factors (TFs) allosterically binding specific genomic sequences, as in procaryotes, but involves epigenetic cues. A challenge is now to provide mechanistic explanation of epigenetic regulation of gene expression. It is related to the fundamental problem of understanding how the eukaryotic cell can respond differently to the same environmental factor according to its type, i.e. the tissue to which it belongs. Very few papers mention the relationship between epigenetic cues, structural features of the chromatin fiber, and transcriptional regulation, see [72]. We here explicitly formulate this relationship within the unifying framework of chromatin allostery, focusing on the mechanistic analysis of the allosteric effect, namely the structural coordination between the triggering event and the distant regulatory event .

In the next Section 2, we briefly review the notion DNA allostery and in Section 3, we discuss how its modalities and functional consequences may be widely enlarged when DNA is embedded in a compact chromatin fiber. In Section 4, we argue that understanding epigenetic control of gene expression is a challenge to the concept of allostery. We propose a scenario in which allosteric transitions of the chromatin fiber induced by histone acetylation trigger the binding of TFs, themselves allosterically prone to binding, in a multiscale way. In Section 5, we discuss how epigenetically controlled shifts between accessible chromatin fiber conformations may control the assembly of DNA-bound regulatory complexes, introducing a notion of large-scale and statistical chromatin allostery. In Section 6, we argue that to understand what could be an epigenetic code, we have to identify the physical adaptors mediating the arbitrary relationship between the codewords, namely histone-tail covalent modifications, and the transcription-related events they encode. We claim that a good candidate is the chromatin fiber and its allosteric transitions. We conclude in Section 7 by opening an allosteric perspective on the epigenetic control of cell differentiation.

2. DNA allostery

DNA allostery is a natural extension of the notion introduced for proteins: it refers to any mechanism in which some local effector event triggers a more extensive DNA conformational transition, which in turns controls distal DNA-binding events. This mechanism, where the coordination of two protein-binding events occur *through* DNA [31], should not be confused with the interplay of an allosteric transition of a protein and its binding to DNA, investigated in numerous works ambiguously associated with the keywords “DNA” and “allostery” in bibliographic searches. We here recall basic features of DNA allostery, as a basis on which to rely for introducing, in Section 3, a more complex notion of multiscale allostery, in which chromatin-embedded DNA allostery is itself monitored at the chromatin fiber level.

2.1. Protein-binding event may trigger DNA allosteric transitions

It is acknowledged that DNA local structure influences DNA-protein binding affinity [7]. This fact is currently invoked in the recruitment of DNA-binding proteins and recognition of their binding sites. It is less acknowledged that this modulation also provides a strong basis for applying the concept of allostery to DNA [12]. As soon as DNA stiffness and structural cohesiveness extend beyond a single binding site, the strain induced by a first binding event generates a DNA deformation over a region embedding other binding sites, and possibly modifies their binding affinity. This is typically the case when the DNA deformation anticipates the DNA structure in the complex formed by the binding site and its ligand. And pertinent to eukaryotic gene expression, closely positioned nucleosomes have been shown quite recently to modify the binding affinity of a transcription factor to a response element inserted in a neighboring linker DNA [31]. Note that speaking of DNA deformation implicitly refers to a lock-and-key mechanism, which may not be the case: an alternative view is a shift of an equilibrium between two pre-existing DNA conformations upon the first binding event (conformational capture). The debate between the scenarios of Kohsland-Némethy-Filmer [32] and Monod-Wyman-Changeux [43] is thus revived in the context of DNA. In the terminology of allostery, the first factor binding DNA acts as an effector, controlling through an extended DNA conformational change the processes occurring at distant sites. We here recover a specific feature of allostery, namely action at a distance, beyond the range of molecular interactions and recognition.

2.2. Effectors may not be proteins

Effectors in DNA allostery are not necessarily a protein. A first example is provided by intercalators, e.g. ethidium bromide. These molecules are efficient effectors capable for instance to convert a left-handed Z-DNA to a right-handed form which promotes further bindings [10]. We proposed that similarly, DNA buckling induced by a first intercalation within a DNA stretch mechanically constrained at both extremities, allows further intercalations, otherwise energetically prohibited, see Fig. 1A, [69]. Cooperative assembly of complexes, possibly favored by DNA allostery, have also been reported [44]. Other non-B DNA forms have recently attracted much attention, as potential mediators of transcriptional control, using as allosteric effectors drugs inducing transitions to such forms [20, 56]. Another example, involving a non-proteic effector, is given by polyamide binding in the minor groove of DNA and modifying the TF/DNA binding interfaces hence affinities [14]. A remarkable experiment has been the design of an artificial effector (a polyamide hairpin), whose binding mimics the conformational change in DNA major groove induced by the first protein binding in DNA minor groove and ensuing protein binding facilitation [45]. As polyamide has no possible direct interactions with the second protein binder, this experiment delineated the contribution of DNA conformational changes, ruled out an explanation of binding cooperativity based on a direct interaction between binding proteins, and thus demonstrated the allosteric nature of the facilitation. Further experimental check has been the observation of the crystal structure of the complex, directly evidencing the structural change experienced by DNA upon binding the allosteric effector [45].

2.3. Allosteric transitions occur in mechanically constrained DNA

The conditions for DNA allostery involve mechanical stiffness, propagating the local structural change induced by the first binding event over a longer DNA stretch. Such mechanical constraints preventing the relaxation of the binding-induced local stresses are observed *in vivo*, typically due to supercoiling in bacterial plasmid [62], or to embedding in a condensed chromatin fiber in eukaryotic organisms [33]. Allosteric modulation of DNA controlling protein binding deserves acknowledgements, since it questions the affinity values and scenarios established for naked DNA in a test-tube, and may induce a long-range coordination between the binding events.

Note that the focus in DNA allostery studies is often on thermodynamical coupling of binding events and associated cooperativity [12, 45]. We also promote a mechanistic explanation in terms of a structural long-range coupling; the former is well-suited to fit experimental, e.g. calorimetric, data but the latter gives more insights on the structural, mechanical or topological determinants [18, 36].

2.4. DNA transitions are not all allosteric transitions

Allosteric transitions, as described above, are only a class of the various structural changes that can be observed in DNA, characterized by the structural coordination of distant events along DNA. In particular, they should be distinguished from salt-induced shape transitions of DNA or chromatin fiber. The latter are phase transitions induced by electrostatic effects; the control parameter is the concentration of surrounding counter-ions or salt [9]. The degree of cooperativity of such transitions is conventionally defined as the number N of cooperative units in the macromolecule. Cooperative units are independent domains where all-or-none transition from one state (e.g. coil) to another (e.g. globule) occur [8]. The degree of cooperativity is given by the vant Hoff equation $N = \Delta H^{cal}/\Delta H^{eff}$ where ΔH^{cal} is the enthalpy of conversion of one macromolecule from one state (e.g. coil) to the other (e.g. globule), and ΔH^{eff} the corresponding enthalpy for only one cooperative subunit. Both ΔH^{cal} and ΔH^{eff} can be measured simultaneously by scanning microcalorimetry in a single experiment [7]. This notion of cooperativity differs from that related to homotropic allostery, which is evaluated through Hill exponent involved in the ligand binding kinetics [7, 67]. DNA allosteric transitions actually appear in two specific instances: cases of homotropic allostery, where the DNA conformational transition triggered by the binding of a first factor (e.g. intercalating residue) increases the binding affinity of additional residues of the same type, with no direct interactions between the binding residues [8, 36]; and cases of heterotropic allostery, ensuring signal transmission, through the geometrical, mechanical or topological coordination of a triggering event, typically the binding of an effector onto DNA, and a functional event, typically the binding at a distant DNA site of a different molecule [8].

3. Allosteric mechanisms in chromatin-embedded DNA

3.1. H3-tail acetylation may control linker-DNA allosteric transitions

It has been proposed already long ago that long-range allosteric transitions of duplex DNA triggered by a local structural change may be involved in transcriptional regulation [55]. Allosteric under-winding of DNA may be involved in transcriptional regulation, by controlling the binding affinity of a TF [3]. However, the longest range

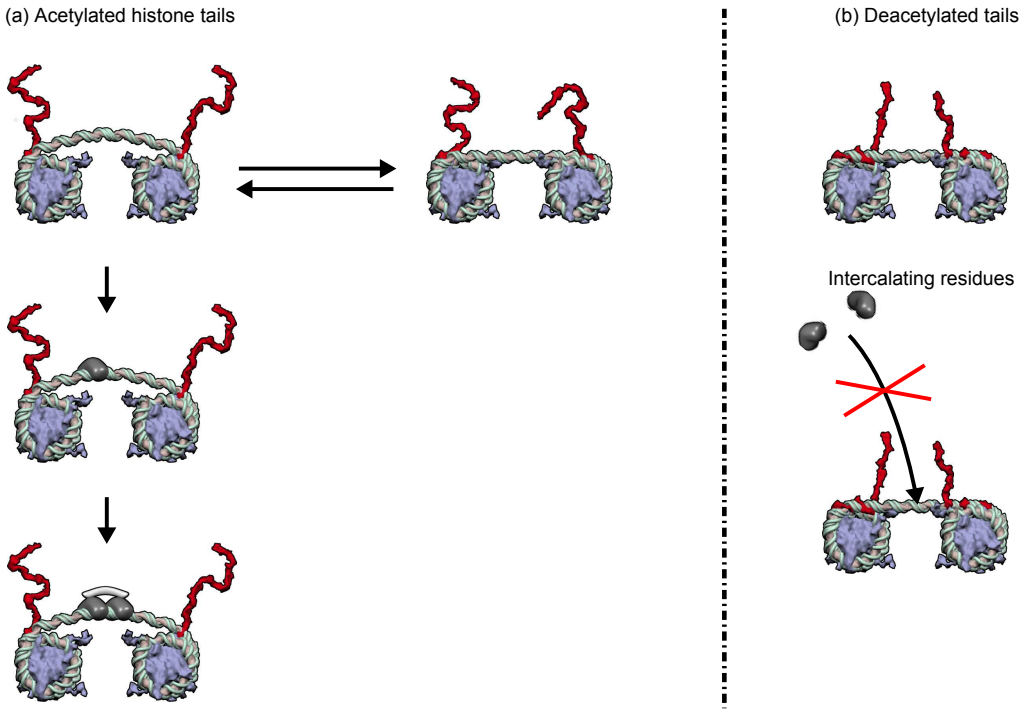


Figure 1. Epigenetically controlled DNA allostery in a chromatin fiber. H3-histone tail acetylation controls the strength of linker DNA binding onto the nucleosomes, sketched as grey disks, whose orientation is fixed by their embedding in the condensed 30-nm chromatin fiber. (Left) When H3 histone tails (in red) are acetylated, DNA anchoring onto nucleosomes is loose and accommodates two linker DNA conformations: a straight T-form (right), where even a single binding is barely possible, and a buckled R-form (left) prone to multiple bindings. The binding of a first intercalating protein residue shifts the conformational equilibrium, in an instance of conformational capture: linker DNA buckling is stabilized and energetically favors subsequent intercalations, in a typical instance of DNA allosteric transition [69]; in particular, the binding of bis-intercalating proteins is possible [36]. (Right) The tight anchoring of linker DNA onto nucleosomes in a chromatin fiber with deacetylated histone tails (in red) prevents DNA buckling transition; multiple intercalations are then energetically prohibited due to mechanical constraints [69].

of such allosteric transitions is only about a hundred of bps, hence DNA allostery alone is likely not to be the only ingredient in transcriptional regulation. Above all, in eukaryote organisms, DNA is embedded within a chromatin fiber, which generates constraints at the DNA level [37]. Chromatin modifications influence the mechanical constraints experienced by linker DNA and in turn monitor the processes occurring at the DNA level, in particular the linker DNA allosteric capabilities, in a multiscale way. Note that the principle of such a multiscale allostery is reminiscent of the global control of the range of allosteric transitions observed in multi-subunit proteins, typically haemoglobin [67]: kinetic studies demonstrate that only a subdomain is

involved in the allosteric behavior, presumably due to the limited cohesiveness of the protein. We henceforth expect a more drastic and more adaptive role in transcriptional regulation of allosteric mechanisms involving both the condensed 30-nm chromatin fiber and the chromatin-embedded DNA. We proposed a first scenario in which DNA allostery is controlled by histone-tail posttranslational modifications at the fiber level (Fig. 1). H3-histone-tail interactions controls linker DNA anchoring onto the flanking nucleosomes, whose orientation is fixed by their embedding into the chromatin fiber and stacking interactions between spatially adjacent nucleosomes (Fig. 2A). H3-tails acetylation, catalyzed by an histone acetyltransferase (HAT), increases the degrees of freedom of embedded linker DNA and makes possible a buckling instability. The first intercalation event induces or stabilizes a DNA allosteric transition from a straight conformation to a buckled one, in which the binding of further intercalating residues is facilitated [69]. Our mechanisitic scenario is further supported by the fact that acetylation increases the α-helical content of the histone tails of the nucleosome [70], which may dramatically favor the unbinding of the histone tail from the DNA.

3.2. Architectural transcription factors facilitate the formation of complex nucleoprotein assemblies

The eukaryotic HMG-box proteins, including in particular HMGB1 and HMGB2, bind DNA non-specifically and induce or stabilize deformed DNA. A major role of these non-sequence specific DNA-binding proteins is to facilitate the formation of complex nucleoprotein assemblies [59]; Figure 4 of [59] is specially informative in this respect. Alternatively, sequence-specific HMG-box transcription factors, as the cell type-specific TFs LEF-1 or SRY, are proteins that contain both an HMG domain which binds to an HMG box, and a sequence specific domain that is targeted to a specific genomic locus. DNA binding of both specific or non-specific HMG-box transcription factors may be regulated by the acetylation state of the nucleosomes that flank the DNA stretch.

4. Chromatin fiber allostery and multiscale control of transcription

4.1. H4-tail acetylation is a physical epigenetic mark that modulates protein-DNA binding affinity

Literature provides several experimental evidences of the correlation between histone acetylation and transcriptional activity. Among epigenetic modifications, H4 tail acetylation stands at a special place. It is not only involved in recruiting specific histone tail-binding complexes, mainly bromodomains, as most epigenetic marks do. It has also a physical action: acetylation neutralizes the positive charges of the lysine residues, which in vitro eventually leads to disruption of compact chromatin fibers [52]. Therefore H4 tail acetylation is believed to activate transcription because it is related to increased DNA local accessibility within chromatin. This view relies on the implicit paradigm that DNA accessibility is the key of transcriptional regulation. We claim that transcriptional regulation is not only a matter of DNA accessibility, which would be modified by chromatin disruption and subsequent nucleosome mobility. It is mainly a matter of affinity, modified by mechanical constraints experienced by DNA as explained in Sec. 2). Indeed most of the genome is actually accessible, in euchromatin as well as in heterochromatin, see [68] and

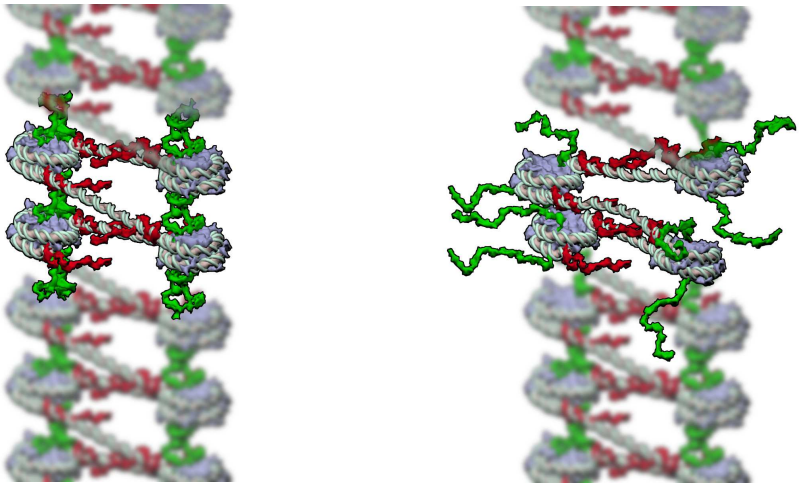


Figure 2. Effect of H4 histone tail acetylation on the chromatin fiber structure. (A) Deacetylated H4-tails (in green) protrude normally to the nucleosome core particle and ensure nucleosome stacking through tail-bridging interactions between their positively charged lysines and the acidic patches of the globular domain of the histone octamer. Deacetylated H3-tails (in red) protrude laterally near the dyad axis at the entry/exit site and constrain the anchoring of the linker DNAs on the histone octamer. (B) Acetylation of H4 histone tails weakens stacking interactions between spatially adjacent nucleosomes and allow unstacked nucleosomes to rotate, hence their flanking linker DNAs to bend despite of the anchoring constraints exerted by the deacetylated H3-tails.

references therein. Even mitotic chromosomes were found to be readily accessible to transcription factor binding [13]. Moreover H4 acetylation does not change the fiber compaction significantly in vitro [2]. We propose instead that H4 tail acetylation modifies at a distance TF binding affinity, by means of a conformational change softening the chromatin fiber by disrupting stacking interactions between spatially adjacent nucleosomes, and in turn modifying the mechanical constraints experienced by embedded DNA. This points to a functional role of the chromatin fiber architecture, and eventually supports the very existence of the fiber itself in vivo and its remarkable conservation during evolution.

4.2. A proof of principle of DNA binding regulation by H4 acetylation

The mechanical and multiscale scenario proposed in [69] and summed up in Fig. 1 involved H3 tail acetylation, which weakens linker DNA anchoring onto the nucleosomes flanking it. We here consider H4-tail acetylation, which weakens the strength of stacking interaction between spatially adjacent nucleosomes (Fig. 2B). Indeed acetylation of core histone tails has distinct effects on nucleosome assemblies depending on whether it affects H3, which modifies lateral interactions, or H4, which modifies stacking interactions. The H4 tail comes out of the nucleosome through the side normal to the nucleosome axis (i.e. normally to the nucleosomal plane) while the H3 tail follows a path between both DNA gyres to the exterior of the nucleosome, hence through the lateral side of it. Therefore the proximal H3 residues, i.e. residues close to the histone fold, exclusively contact DNA stretches comprised in their own nucleosome, unlike H4 which bridges stacked nucleosomes. Interestingly H4K16, which is a proximal residue of the N-term H4 chain, is unique among lysines to make histone-histone contacts between nucleosomes of the same fiber, thus resulting in stacking interactions. Note that these interactions between positively charged lysines are mediated by divalent cations [9].

We propose here an alternative mechanical scenario accounting for TF binding control by H4 acetylation. In this multiscale scenario, sketched in Fig. 3, H4K16 acetylation of the nucleosomes flanking some locus triggers an allosteric transition of the chromatin fiber from a tight and mechanically constrained form to a loose soft form, endowed with more degrees of freedom; in plain terms, the allosteric transition turns a locked form into an unlocked one. While it induces only a fine-tuning of chromatin compaction, this transition has strong consequences in terms of mechanical constraints. Indeed, the fiber softening enables further binding of TFs that bind preferentially to distorted DNA structures, as HMG-box proteins (see Fig. 3). Importantly, allostery is only possible in a compact fiber, experiencing mechanical constraints.

O. Rando recently suggested that histone modifications may not recruit TFs and chromatin regulators to specific loci but may instead allosterically activate their binding [53]. However in this view, allostery is limited to proteins. A typical example is the activation of the Rpd3S deacetylase complex by the H3K36me3 methylation mark. A similar view, albeit less explicit and not mentioning allostery, may be found in [68]. In this EMBO member's review, it is proposed (see figure 4 of [68]) that TFs may be targeted to tissue-specific loci by cooperative protein-protein interactions between matching shapes of the TFs and "niches" in the tissue-specific chromatin type. These "niches" differ from one tissue to another thanks to epigenetic marks. The combinatorial complexity of epigenetic marks and associated chromatin proteins

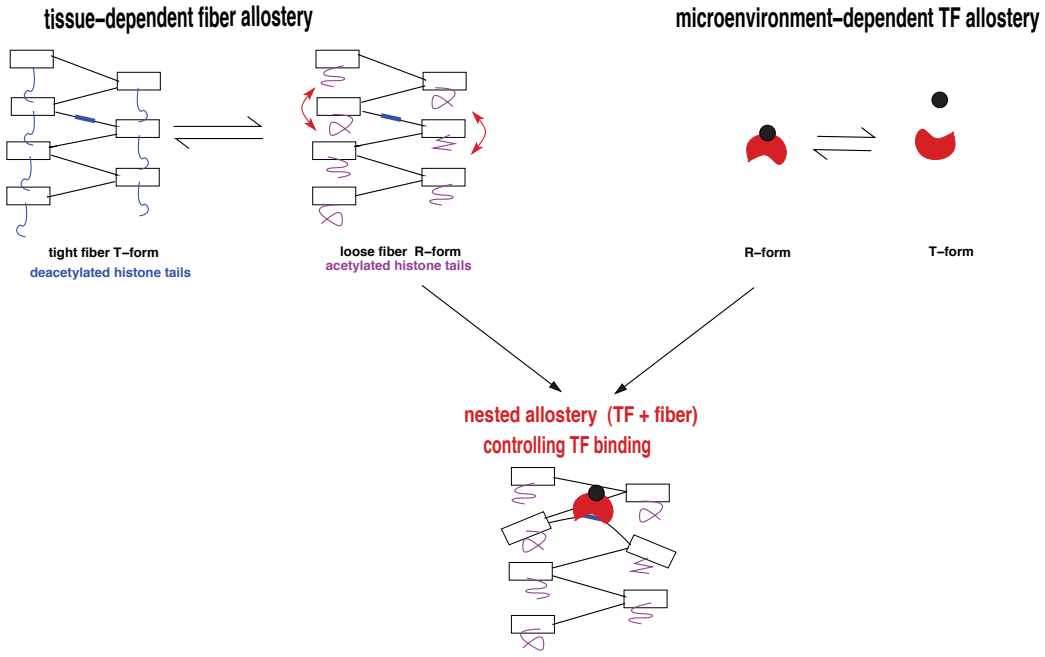


Figure 3. Nested allosteric mechanisms in eukaryote transcriptional regulation. (Top left) Histone acetylation, here H4K16ac triggers an allosteric transition of the fiber from a tight and mechanically constrained form to a loose form, namely from a T-form to R-form when using the terminology of [43]. In this transition, nucleosomes and linkers gain respectively tilting and buckling degrees of freedom. This allosteric mechanism is tissue dependent, having as effector the HATs and as active sites all the potential protein binding sites present in the embedded DNA linkers. (Top right) Molecular effectors brought in the microenvironment through signaling or metabolic pathways trigger an allosteric transition of TFs, from an inactive T-form to an R-form capable to bind to sufficiently relaxed DNA linker. (Bottom) The conjunction of the two allosteric transitions at the TF and fiber levels controls TF binding therefore participates in transcriptional regulation.

has been shown to reduce in *Drosophila* to five main chromatin types [23] and more recently refined to nine states [30]. The above multiscale scenario, Fig. 3, offers a mechanical framework unifying these observations: epigenetic control may rely on an allosteric transition softening the chromatin fiber and modifying the mechanical constraints experienced by embedded DNA, hence DNA/TF binding affinity. A physical mechanism of allosteric transmission provides an alternative to chemical competition to achieve epigenetic regulation.

4.3. Acetylation marks facilitate targeting of TFs to their cognate genomic sequence

One may think to an allosteric TF that comprises on a one side a bromodomain, which recognizes an epigenetic pattern of histone modifications, and on the other side a DNA-binding domain able to bind a specific genomic sequence. Such a TF would

be targeted to its cognate binding site selectively in cells where the binding site is flanked by acetylated nucleosomes. In this scheme, there is no mandatory need for allosteric mechanism, it is just a matter of combinatorial recruitment. Alternatively, the bromodomain could as well play the role of an effector in case its binding to the proper epigenetic pattern would allosterically trigger the binding of the DNA-binding domain to its cognate sequence.

However, in such a scenario, the TF would be hampered in its searching for its cognate sequence by the huge number of histone acetylation patterns present in the nucleus. As a result, such a TF would never find its target in due time [57]. This kinetic argument is all the more relevant as the TF copy number is small. Moreover, in a recent review [53], Rando gathers experimental evidences that epigenetic patterns are rarely colocalized with increased transcriptional activity, although an overall correlation is present.

The scenario we propose here answers these two, logical and experimental, objections: an allosteric transition of the chromatin fiber mediates the non specific recognition of histone modification patterns by a factor embedding a bromodomain and the specific binding, at a distance, of a TF onto its cognate genomic sequence. More specifically, histone tail acetylation results in the softening of the fiber hence in more bendable DNA linkers at proximity, thus facilitating the binding of specific TFs, which otherwise could not bind. Importantly, there is no need for specific recognition of acetylated lysines, e.g. by bromodomains, but instead the TF is repulsed from deacetylated nucleosomes, thus accelerating substantially its search for its cognate sequence.

5. Large-scale chromatin fiber allostery

An important aspect of genomic regulation by means of DNA and chromatin structural transitions is its multiscale nature [8]. We have discussed above the functional consequences of DNA allostery and chromatin-embedded DNA (of range tens of bp), then of modifications of the local architecture and stiffness of the chromatin fiber at ranges of a few kb. Larger-scale conformational changes of the chromatin fiber, at kb or even Mb scales, have been recently assessed experimentally within the so-called topologically-associating domains (TADs) corresponding to compaction units of the chromatin [25]. We here argue that such conformational changes are the basis of a large-scale allosteric control of transcriptional regulation.

5.1. Chromatin conformational subsets influence regulatory complex assembly

Actually it has been recently demonstrated in mouse cells that chromatin fiber *in vivo* may adopt many conformations [25]. What is determined is only the distribution of these conformations across time and cells in a population. On the other hand, conformation of the chromatin fiber controls the assembly of complexes of various factors, including TFs [28]. Any change affecting the accessible conformations will thus modify and potentially regulate which complexes will assemble onto DNA, as sketched in Fig. 4. In particular epigenetic modifications may control the subset of accessible chromatin fiber conformations, in a large-scale instance of chromatin allostery. Here the allosteric transition is a shift [18] between the statistically dominant subsets of chromatin fiber conformations, triggered by some effector event, among which we suggest epigenetic marks may play a key role [8]. In this view, epigenetic constraints

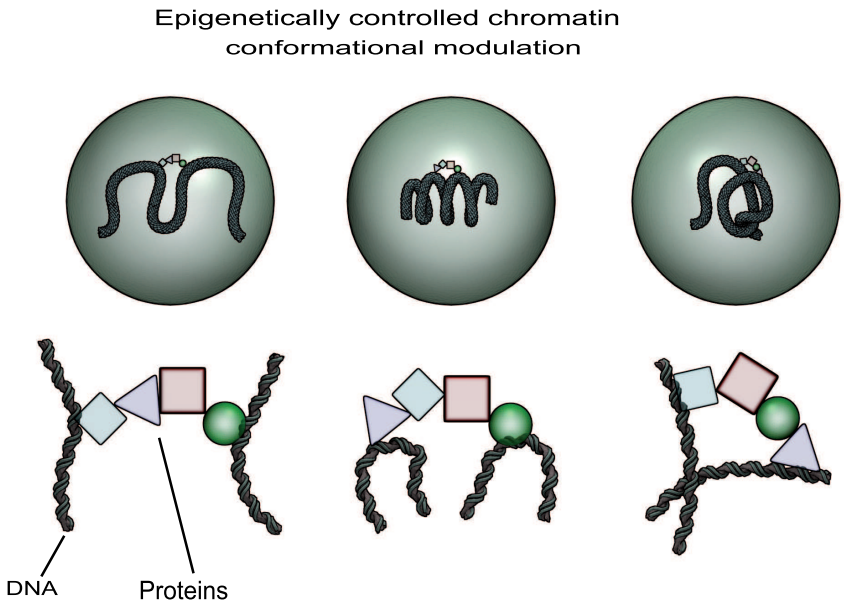


Figure 4. Chromatin conformation controls complex assembly. Physiological factors and signals, possibly epigenetically mediated, control the subset of statistically dominant chromatin fiber conformations (top, each circle correspond to a different cell state). In turn, the fiber conformation controls binding of factors to DNA and subsequent cofactor binding. Shift in the distribution of chromatin fiber conformations channels the assembly of different regulatory complexes.

presumably delineate different subsets of accessible chromatin fiber conformations, which each favors the assembly of dedicated regulatory complexes. Conversely, complexes may stabilize the adapted chromatin conformations, in a large-scale instance of conformational capture. Different complexes associations accommodate different allosteric signal propagation pathways [63]. These physical mechanisms provide a physical alternative to chemical recognition of epigenetic marks. In this way, chromatin fiber allostery may thus play a key role in bridging signal transduction with complexes assembly onto DNA, thus mediating the epigenetic regulation of gene expression.

5.2. Experimental validation

Large-scale chromatin allostery may play an important role in setting the activation of gene transcription by remote enhancer sequences [8]. In mammals, one locus where such a mechanism may be involved is the paradigmatic *Igf2/H19* gene locus (Fig. 5A, map of the mouse locus). It was recently shown that, in the absence of strong long-

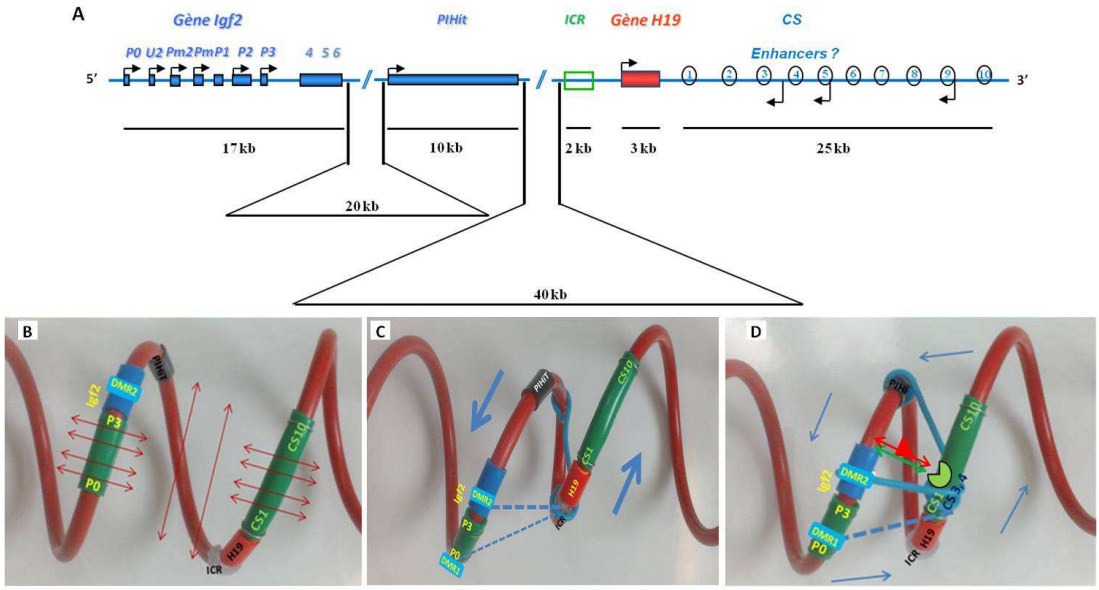


Figure 5. Chromatin fiber allostery and *Igf2* gene remote activation **A**. Linear map of the mouse *Igf2/H19* gene locus. Promoters of the *Igf2* gene (P0-P3) are indicated by black arrows. The locus is containing two other genes named *PIHit* and *H19*. Remote enhancers named CS1-CS10 are located downstream of the *H19* gene and about 90kb away from the *Igf2* gene. The Imprinting-Control Region (ICR), that inherits DNA methylation when paternally inherited, is located 3kb upstream of the *H19* gene. **B**. Folding of the *Igf2/H19* locus in the context of the statistical helix. Note: Even in the absence of locus specific interactions, the *Igf2* promoters (P0-P3) are located close to the enhancers (CS1-CS10) in the tridimensional space. During mouse development, allele specific interactions occur, leading to different folding of the locus on the two parental chromosomes. **C**. On the maternally inherited chromosome, the CTCF (CCCTC-binding Factor) insulator protein binds to the umethylated Imprinting-Control Region (ICR) thus promoting specific interactions with other regions of the locus (blue bars) that move the enhancers away from the *Igf2* promoters (blue arrows). The *Igf2* gene is thus transcriptionally silent on that chromosome. **D** On the paternally inherited chromosome, DNA methylation of the ICR prevents CTCF binding and tissue-specific TFs bound to the enhancers (red triangle, green circle) thus inducing locus specific interactions (blue bars) and stabilizing the contacts between a specific enhancer, here CS4, and the *Igf2* gene. Altogether, at this locus, specific interactions alter the statistical shape of the chromatin fiber in chromosome- and tissue-specific ways, depending on the DNA methylation status of the ICR and on the TFs bound on the enhancers. Such chromatin allosteric effects allow the harmonious regulation of *Igf2* transcription (figure adapted from [61]).

range locus-specific interactions, the mammalian chromatin tends to adopt statistically a helix shape [16]. If we depict the mouse *Igf2/H19* locus in the context of the statistical helix, Fig. 5B, we observe that sets of remote enhancers, known to lie more than 90kb away from the *Igf2* gene, are in fact positioned close to this gene when considered in the tridimensional space. Chromosome Conformation Capture (3C) experiments have shown that, during mammalian development, different specific interactions occur on the two parental chromosomes, leading to different folding of this locus and thus to allele specific expression of the *Igf2* and *H19* genes, what is known as parental imprinting. The differential folding of this locus is driven by the DNA methylation status of a sequence called the Imprinting-Control Region (ICR) [60], which is different on the two parental alleles. On the maternally inherited chromosome, Fig. 5C, the ICR is unmethylated allowing the CTCF (CCCTC-binding Factor) insulator protein to bind [26]. The ICR-CTCF insulator then interacts with several other sequences of the locus [34] leading the enhancers to move away from the *Igf2* promoters that are thus silenced on this chromosome. On the paternally inherited chromosome, DNA methylation of the ICR prevents CTCF binding and the interactions between enhancers and the *Igf2* gene take place at this locus in a tissue-dependent manner, Fig. 5D [17, 22]. It is thus tempting to speculate that, on that chromosome, fixation of tissue-specific TFs on the enhancers is inducing large-scale chromatin allosteric effects that help to fine-tune the tridimensional organization of the locus, favoring in this way the interaction of specific enhancers with the *Igf2* gene for appropriate, tissue-specific, transcriptional activation.

5.3. Large-scale supercoiling

What kind of physical constraints are involved in the epigenetic selection of conformational subsets? Topological constraints, in particular large-scale supercoiling, have ben shown to play a major role. So, eukaryotic chromosomes are organised into under-wound and over-wound supercoiling domains (about 100kb), whose boundaries are formed and maintained by CTCF binding sites [24]. In particular, the boundaries of TADs which are also enriched for CTCF binding sites, act as supercoiling boundary elements. These supercoiling domains are formed through polymerase and topoisomerase activity. For instance, we recently shown that transcription elongation is theoretically possible through a compact fiber thanks to a coordinate modification of the nucleosomes allosterically triggered by the processing polymerase [5, 36]; indeed, the torsional stress generated by the active polymerase is sufficient to promote the sequential transition of stacked nucleosomes to a permissive state, the so-called reversome, allowing the passage of the polymerase and the processivity of the transcription. This wave of positive supercoiling induces in turn a transcription coupled allosteric underwinding several kb *upstream* of the polymerase [33]. We finally wish to suggest here that the statistical helix conformation of the mouse *Igf2/H19* gene locus may result from supercoiling constraints. Moreover, as shown above, imprinting dependent contacts may locally tune supercoiling constraints to allosterically favor remote specific contacts as far as 100kb apart.

6. A new perspective on the epigenetic code

6.1. Histone code is a debated notion

The discovery of the wealth of histone-tail covalent modifications raised a challenging issue: understand their role, if any, in the regulation of transcription. It has been proposed that these modifications and the associated combinatorics constitute a code, the histone code [29]. Struggle for deciphering this code was accompanied by a debate on whether it is actually a code [27, 35, 64]. It has been underlined that the term "histone code" covers several themes, either the combinatorics of histone modifications, or their heritability, or their molecular recognition by binding partners [53], which do not directly pertain to the code issue. Alternative names were proposed, in particular the epigenetic code [65] or the chromatin code [6]. This code is usually described as a combinatoric means for generating specific recognition interfaces on the nucleosome [66] and recruiting specific cofactors. Our proposed notion of chromatin allostery triggered by histone-tail modifications offers a different, mechanical view on this code issue and opens a novel research direction. In the present paper, we have illustrated this scenario with the emblematic examples of H3-tail and H4-tail acetylation, however its principle applies to any other histone modification. Numerous other potential situations can be found in the literature, that would be deserved to be reconsidered and investigated in this perspective, e.g. the effect of phosphorylation and ubiquitination of core histones[72], or the importance of chromatin fiber conformational fluctuations in X chromosome inactivation [25].

6.2. A biological code relies on a co-evolved adaptor

To settle the debate, we need to briefly state what is a code in a biological context. The reference is obviously the genetic code. We will consider that a correspondence between a set of codewords (codons, histone modifications, ...) and a set of objects or events (amino-acids, transcriptional activity, ...) deserves to be termed a code if it is a gratuitous correspondence mediated by a co-evolved adaptor. In the case of the genetic code, the adaptor is a charged tRNA, whose anticodon loop recognizes the codon, while the charged stem brings in the cognate amino-acid. It should be possible to change a code without changing the laws of physicochemistry, by designing another adaptor [35]; this is the case for the genetic code, where artificial code variants have been designed by modifying the tRNAs, thus demonstrating the arbitrariness of the code [71].

We argue that an allosteric entity fulfills this requirement of arbitrariness of a code and can be seen as an adaptor in the above sense. The arbitrariness, also termed gratuity, of a code comes from the evolutionary design of its adaptors. The gratuity of an allosteric relationship has been recognized long ago [40, 41]. Quoting [41]: *"The absence of any inherent obligatory chemical analogy or reactivity between substrate and allosteric effector appears to be a fact of extreme biological importance."* This gratuity is related to the evolutionary origin of allostery , which has been experimentally demonstrated for proteins by means of phylogenetic analyses or directed evolution experiments [49]. The involvement of natural selection in the design of allosteric entities and mechanisms was already underlined in [41]: *" The specific structure of any enzyme-protein is of course a pure product of selection, necessarily limited, however, by the structure and chemical properties of the actual reactants."* Such statements equally apply in the context of DNA and chromatin fiber allostery [36].

6.3. Allosteric signaling can be seen as a one-word code

The notion of code is besides indissociable from a similarly debated notion, that of a "signal". It was already underlined in [41] that allosteric effectors should be seen as signals, or codewords in the perspective of a code. The frontier between a signal and a code is besides fuzzy. Indeed, the functioning of an allosteric enzyme can yet be seen as a code, with a single codeword, namely the effector, encoding a metabolic or hormonal signal; the adaptor in this degenerate code is the allosteric enzyme, "translating" the codeword into an actual chemical event at the active site. This encoding has been achieved by evolution through natural selection. In this respect, any allosteric relationship between an effector or a set of effectors, and a regulatory mechanism as some active sites has the semiotic properties required to be termed a code [4]. The decoding of such a code is precisely achieved by the allosteric entity [47]. Cellular signaling networks also rely on allosteric entities (most often proteins) and allosteric principles, in relating pathways with no chemical necessity [46]. An example of allosteric signaling modules is provided by G-proteins and G-protein-coupled receptors [73]; the latter are transmembrane receptors that experience an allosteric conformational change in response to signals coming from outside the cell, which in turn triggers an activating allosteric transition of intracellular G-proteins bound to them. The parallel established between chromatin fiber and a signaling network [54] simply stems from the fact that both necessarily involve an allosteric step. Indeed, according to the very definition of a signal [58], there is no physicochemical necessity between the input signal and the response it is triggers.

6.4. Chromatin fiber allostery may be involved in the translation of the histone code into transcription regulatory events

We have here proposed that chromatin fiber allosteric transitions may play a key role in decoding the histone code into transcriptional regulation. Histone modifications allosterically modify TF structure and their binding affinities/capabilities [53]. We believe that only the involvement of such an allosteric mechanism justifies to speak of a "chromatin code". Allostery is indeed an example of gratuity [40]. The evolutionary origin of allostery [49] supports the bona fide nature of this code, allowing the implementation of information transfer along the genome. We suggest that the histone code would be better termed a context-dependent transcription regulatory code [54], translated by the chromatin fiber allosteric transitions. Context-dependence arises through the involvement of the chromatin fiber superstructure. Quoting [41], this context-dependence can only be achieved through an allosteric mechanism: *"In fact it seems difficult to imagine any biochemical mechanism other than allosteric which could allow a single chemical signal to be understood and interpreted simultaneously in different ways by entirely different systems."*

7. Allostery: a wide-range unifying concept

The notion of allostery provides a way to account for the evolutionary imprint in the functional activity of protein or nucleic acids. We proposed that the concept of allostery may also apply to genomic functions, in particular transcriptional regulation. The analog of end-product inhibition would be a protein allosterically regulating its own transcription. However, the logical scheme in gene regulation is in general more

complex than a mere feedback loop. It takes the form of regulatory networks, in which a synthesized protein allosterically regulates the synthesis of another one, at the level of the transcription of its genetic sequence.

Interplay between biochemical regulation and mechanics gives a unique functional role to the 30-nm chromatin fiber, and provides a target for epigenetic marks, in particular posttranslational modifications of histone tails. It moreover provides a general mechanism underlying transcription within facultative heterochromatin. Indeed, transcription initiation may not require chromatin decondensation, but simply the softening of the fiber, either through H3 tail acetylation (Fig. 1) or through H4K16 acetylation (Fig. 2). Fiber softening mainly expresses in the alleviation of mechanical stress experienced by linker DNA and in particular the weakening of linker anchoring onto the nucleosomes flanking it, due to the modification of H3 tails (Fig. 1) and in a weakening of the stacking interactions between spatially adjacent nucleosomes in the fiber, due to the modification of H4 tails (Fig. 2B). Moreover, it has been observed that some transcriptional activators have an acetyltransferase activity or interact with an histone acetyltransferase, and that some transcriptional repressors act by recruiting a deacetylase [72]. This softening may eventually enable TBP binding to the TATA-box and further assembly of TFIID, even in a compact fiber [36]. It is to note that our allosteric scenario takes place in the condensed 30-nm chromatin fiber (as opposed to the 10-nm bead-on-strings nucleofilament). It intends to explain the very first step of gene activation while nucleosomes are still present, namely the fiber opening which allows the recruitment of remodeling factors, and the subsequent displacement of nucleosomes. It has indeed been experimentally established that the nucleosome depletion at TATA-boxes occur after chromatin fiber opening and remodeling [1, 15, 38, 39].

The notion of chromatin fiber allostery offers a plausible explanation of how the cell performs signal transduction and signal integration in the genome function, in a spatially differentiated way. This potentiality of chromatin fiber structure and superstructure may explain cell differentiation, a goal at which ultimately aimed the work of Monod, Changeux and Jacob. It indeed provides a pathway towards different gene expression in cells of a given lineage, shifting the main puzzle about cell differentiation to the stabilization of definite cell types. In this respect it is worth underlining that only eukaryotic species exhibit a stable pluricellular structure. Prokaryotic species are known to exhibit collective behavior and changes in cell phenotypes (colonies of bacteria, *Dictyostelium D.* slime mould). But these changes are transiently induced by direct environmental constraints and cell signaling. Although the underlying logic might be the same and the discrepancy lay only in the persistence times of the collectively organized structures, it is tempting to locate the difference in the chromatin, which is one of the most conserved entities in the eukaryotic realm, and in its functional role in transcription.

References

[1] Agalioti T, Lomvardas S, Parekh B, Yie J, Maniatis T & Thanos D 2000 Cell **103** 931–7
[2] Allahverdi A, Yang R, Korolev N, Fan Y, Davey CA, Liu CF & Nordenskiöld L 2011 Nucl. Acids Res. **39** 1680–91
[3] Ansari AZ, Chael ML & Ohalloran TV 1992 Nature **355** 87–9
[4] Barbieri M 2003 The Organic codes – an introduction to semantic biology (Cambridge: Cambridge University Press)
[5] Bécavin C, Barbi M, Victor JM & Lesne A 2010 Biophys. J. **98**, 824–33.

[6] Benecke A 2006 Eur. Phys. J. E **19** 379–84
[7] Cantor CR & Schimmel PR 1980 Biophysical chemistry: Part III: the behavior of biological macromolecules (San Francisco: Freeman)
[8] Caré B, Carrivain P, Forné T, Victor JM & Lesne A 2014 Commun. Theor. Phys. (in press)
[9] Carrivain P, Cournac A, Lavelle C, Lesne A, Mozziconacci J, Paillusson F, Signon L, Victor J M & Barbi M 2012 Soft Matter **8** 9285–301
[10] Chaires JB 1986 J. Biol. Chem. **261** 8899–907
[11] Chaires JB 1997 Biopolymers **44** 201–14
[12] Chaires JB 2008 ACS Chem. Biol. **3** 207–9
[13] Chen D, Dundr M, Wang C, Leung A, Lamond A, Misteli T & Huang S 2005 J. Cell Biol. **168** 41–54
[14] Chenoweth DM & Dervan PB 2009 Proc. Natl. Acad. Sci. USA **106** 13175–9
[15] Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M & Zaret KS 2002 Mol. Cell **9** 279–89
[16] Court F, Miro J, Braem C, Lelay-Taha MN, Brisebarre A, Atger F, Gostan T, Weber M, Cathala G & Forné T 2011 Genome Biol. **12** R42.
[17] Court F, Baniol M, Hagège H, Petit JS, Lelay-Taha MN, Carbonell F, Weber M, Cathala G & Forné T 2011 Nucl. Acids Res. **39** 5893–906
[18] Cui Q & Karplus M 2008 Protein Sci. **17** 1295–307
[19] Dion M, Altschuler S, Wu L & Rando OJ 2005 Proc. Natl. Acad. Sci. USA **102** 5501–6
[20] Du X, Wojtowicz D, Bowers AA, Levens D, Benham CJ & Przytycka TM 2013 Nucl. Acids Res. **41** 5965–77
[21] Edelstein SA (ed.) J. Mol. Biol. **425**, issues (9) and (13) on "Allosteric interactions and biological regulation"
[22] Engel N, Raval AK, Thorvaldsen JL & Bartolomei SM 2008 Hum. Mol. Gen. **17** 3021–9
[23] Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ & van Steensel B 2010 Cell **143** 212–24
[24] Naughton C, Avlonitis N, Corless S, Prendergast J.G, Mati I.K, Eijk P.P, Cockroft S.L, Bradley M, Ylstra B & Gilbert N 2013 Nat. Struct. Mol. Biol. **20** 387–395
[25] Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, Tiana G & Heard E 2014 Cell **157** 950–63
[26] Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM & Tilghman SM 2000 Nature **405** 486–9
[27] Henikoff S 2006 The Scientist **20** 38–9
[28] Ho Y, Shewchuk BM, Liebhaber SA & Cooke NE 2013 Mol. Cell. Biol. **33** 1723
[29] Jenuwein T & Allis CD 2001 Science **293** 1074–80
[30] Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, Linder-Basso D, Plachetka A, Shanower G, Tolstorukov MY, Luquette LJ, Xi R, Jung YL, Park RW, Bishop EP, Canfield TK, Sandstrom R, Thurman RE, MacAlpine DM, Stamatoyannopoulos JA, Kellis M, Elgin SC, Kuroda MI, Pirrotta V, Karpen GH & Park PJ 2011 Nature **471** 480–5
[31] Kim S, Broströmer E, Xing D, Jin J, Chong S, Ge H, Wang S, Gu C, Yang L, Gao YQ, Su XD, Sun Y & Xie XS 2013 Science **339** 816–9
[32] Koshland DE Jr, Némethy G & Filmer D 1966 Biochemistry **5** 365–8
[33] Kouzine F, Sanford S, Elisha-Feil Z & Levens D 2008 Nat. Struct. Mol. Biol. **15** 146–54
[34] Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W & Ohlsson R 2006 Proc. Natl. Acad. Sci. USA **103**10684–9
[35] Lesne A 2006 Eur. Phys. J. E **19** 375–7
[36] Lesne A, Bécavin C & Victor JM 2012 Phys. Biol. **9** 013001
[37] Lesne A & Victor JM 2006 Eur. Phys. J. E **19** 279–90
[38] Lomvardas S & Thanos D 2001 Cell **106** 685–96
[39] Lomvardas S & Thanos D 2002 Mol. Cell **9** 209–11
[40] Monod J 1972 Chance and necessity. An essay on the natural philosophy of modern biology (London: Collins)
[41] Monod J, Changeux JP & Jacob F 1963 J. Mol. Biol. **6** 306–29
[42] Monod J & Jacob F 1961 Cold Spring Harb. Symp. Quant. Biol. **26** 389–401
[43] Monod J, Wyman J & Changeux JP 1965 J. Mol. Biol. **12** 88–118
[44] Moretti R & Ansari AZ 2008 Biochimie **90** 1015–25
[45] Moretti R, Donato LJ, Bresinski ML, Stafford RL, Hoff H, Thorson JS, Dervan PB & Ansari AZ 2008 ACS Chem. Biol. **3** 220–9
[46] Nussinov R 2013 Phys. Biol. **10** 045004
[47] Nussinov R, Ma B, Tsai CJ & Csermely P 2013 Structure **21** 1509–21

[48] Nussinov R, Tsai CJ & Ma B 2013 Annu. Rev. Biophys. **42** 169–89
[49] Peracchi A & Mozzarelli A 2010 Biochim. Biophys. Acta — Proteins and Proteomics **1814** 922–33
[50] Perutz MF 1962 X-ray analysis of haemoglobin, Nobel Lecture, December 11, 1962.
[51] Perutz MF 1989 Q. Rev. Biophys. **22** 139–6
[52] Potoyan DA & Papoian GA 2012 Proc. Natl. Acad. Sci. USA **109** 17857–62
[53] Rando OJ 2012 Curr. Op. Gen. Dev. **22** 148–55
[54] Schreiber L & Bernstein BE 2002 Cell **111** 771–8
[55] Schurr JM, Delrow JJ, Fujimoto BS & Benight AS 1997 Biopolymers **44** 283–308
[56] Sharma S 2001 J. Nucl. Acids **2011** 724215
[57] Sheinman M, Bénichou O, Kafri Y & Voituriez R 2012 Rep. Prog. Phys. **75** 026601
[58] Stock AM, Robinson VL & Goudreau PN 2000 Annu. Rev. Biochem. **69** 183–215
[59] Thomas JO & Travers AA 2001 Trends in Biochem. Sci. **26** 167–74
[60] Thorvaldsen JL, Duran KL & Bartolomei MS 1998 Genes & Dev. **12** 3693–702
[61] Tran VG 2014 *Régulation de lexpression du gene Igf2 : nouveaux promoteurs et implication de longs ARN non-codants*, PhD Thesis, University Montpellier II
[62] Travers A & Muskhelishvili G 2005 Curr. Opin. Genet. Dev. **15** 507–14
[63] Tsai CJ & Nussinov R 2011 Biochem. J. **439** 15-25
[64] Turner BM 2006 The Scientist **20** 38–9
[65] Turner BM 2007 Nat. Cell Biol. **9** 2-6
[66] Turner BM 2012 Trends Gen. **28** 436–45
[67] van Holde KE, Miller KI & van Olden, E 2000 Biophys. Chem. **86**, 165-72.
[68] Van Steensel B 2011 EMBO J **30** 1885–95
[69] Victor JM, Ben-Haïm E & Lesne A 2002 Phys. Rev. E **66** 060901
[70] Wang X, Moore SC, Laszckzak M, Ausió J 2000 J. Biol. Chem. **275** 35013–20.
[71] Wang L & Schultz PG 2005 Angew. Chem. Int. Ed. **44** 34–66
[72] Wolffe AP & Guschin D 2000 J. Struct. Biol. **129** 102–22
[73] The studies of G-proteins-coupled receptors, The Nobel Prize in Chemistry 2012 - Advanced Information, 26 Oct 2012.
`http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2012/advanced.html`

Contents

1 Introduction **2**

2 DNA allostery **3**
2.1 Protein-binding event may trigger DNA allosteric transitions 4
2.2 Effectors may not be proteins 4
2.3 Allosteric transitions occur in mechanically constrained DNA 5
2.4 DNA transitions are not all allosteric transitions 5

3 Allosteric mechanisms in chromatin-embedded DNA **5**
3.1 H3-tail acetylation may control linker-DNA allosteric transitions . . . 5
3.2 Architectural transcription factors facilitate the formation of complex nucleoprotein assemblies 7

4 Chromatin fiber allostery and multiscale control of transcription **7**
4.1 H4-tail acetylation is a physical epigenetic mark that modulates protein-DNA binding affinity 7
4.2 A proof of principle of DNA binding regulation by H4 acetylation . . . 9
4.3 Acetylation marks facilitate targeting of TFs to their cognate genomic sequence 10

5 Large-scale chromatin fiber allostery **11**
5.1 Chromatin conformational subsets influence regulatory complex assembly 11
5.2 Experimental validation 12
5.3 Large-scale supercoiling 14

6 A new perspective on the epigenetic code **15**
6.1 Histone code is a debated notion 15
6.2 A biological code relies on a co-evolved adaptor 15
6.3 Allosteric signaling can be seen as a one-word code 16
6.4 Chromatin fiber allostery may be involved in the translation of the histone code into transcription regulatory events 16

7 Allostery: a wide-range unifying concept **16**

Figure captions

Figure 1: Epigenetically controlled DNA allostery in a chromatin fiber. H3-histone tail acetylation controls the strength of linker DNA binding onto the nucleosomes, sketched as grey disks. (Left) When H3 histone tails (in red) are acetylated, DNA anchoring onto nucleosomes is loose and accommodates two linker DNA conformations: a straight T-form (right), where even a single binding is barely possible, and a buckled R-form (left) prone to multiple bindings. The binding of a first intercalating protein residue shifts the conformational equilibrium, in an instance of conformational capture: linker DNA buckling is stabilized and energetically favors subsequent intercalations, in a typical instance of DNA allosteric transition [69]; in particular, the binding of bis-intercalating proteins is possible [36]. (Right) The tight anchoring of linker DNA onto nucleosomes in a chromatin fiber with deacetylated histone tails (in red) prevents DNA buckling transition; multiple intercalations are then energetically prohibited due to mechanical constraints [69].

Figure 2: effect of H4-tail acetylation on the chromatin fiber structure. (A) Deacetylated H4-tails (in green) protrude normally to the nucleosome core particle and ensure nucleosome stacking through tail-bridging interactions between their positively charged lysines and the acidic patches of the globular domain of the histone octamer. Deacetylated H3-tails (in red) protrude laterally near the dyad axis at the entry/exit site and constrain the anchoring of the linker DNAs on the histone octamer. (B) Acetylation of H4 histone tails weakens stacking interactions between spatially adjacent nucleosomes and allow unstacked nucleosomes to rotate, hence their flanking linker DNAs to bend despite of the anchoring constraints exerted by the deacetylated H3-tails.

Figure 3: Nested allosteric mechanisms in eukaryote transcriptional regulation. (Top left) Histone acetylation, here H4K16ac triggers an allosteric transition of the fiber from a tight and mechanically constrained form to a loose form, namely from a T-form to R-form when using the terminology of [43]. In this transition, nucleosomes and linkers gain respectively tilting and buckling degrees of freedom. This allosteric mechanism is tissue dependent, having as effector the HATs and as active sites all the potential protein binding sites present in the embedded DNA linkers. (Top right) Molecular effectors brought in the microenvironment through signaling or metabolic pathways trigger an allosteric transition of TFs, from an inactive T-form to an R-form capable to bind to sufficiently relaxed DNA linker. (Bottom) The conjunction of the two allosteric transitions at the TF and fiber levels controls TF binding therefore participates in transcriptional regulation.

Figure 4: Chromatin conformation controls complex assembly. Physiological factors and signals, possibly epigenetically mediated, control the subset of statistically dominant chromatin fiber conformations (top, each circle correspond to a different cell state). In turn, the fiber conformation controls binding of factors to DNA and subsequent cofactor binding. Shift in the distribution of chromatin fiber conformations channels the assembly of different regulatory complexes, in a large-scale instance of chromatin allostery.

Figure 5: Chromatin fiber allostery and *Igf2* gene remote activation A.

Linear map of the mouse *Igf2/H19* gene locus. Promoters of the *Igf2* gene (P0-P3) are indicated by black arrows. The locus is containing two other genes named *PIHit* and *H19*. Remote enhancers named CS1-CS10 are located downstream of the *H19* gene and about 90kb away from the *Igf2* gene. The Imprinting-Control Region (ICR), that inherits DNA methylation when paternally inherited, is located 3kb upstream of the *H19* gene. **B**. Folding of the *Igf2/H19* locus in the context of the statistical helix. Note: Even in the absence of locus specific interactions, the *Igf2* promoters (P0-P3) are located close to the enhancers (CS1-CS10) in the tridimensional space. During mouse development, allele specific interactions occur, leading to different folding of the locus on the two parental chromosomes. **C**. On the maternally inherited chromosome, the CTCF (CCCTC-binding Factor) insulator protein binds to the umethylated Imprinting-Control Region (ICR) thus promoting specific interactions with other regions of the locus (blue bars) that move the enhancers away from the *Igf2* promoters (blue arrows). The *Igf2* gene is thus transcriptionally silent on that chromosome. **D** On the paternally inherited chromosome, DNA methylation of the ICR prevents CTCF binding and tissue-specific TFs bound to the enhancers (red triangle, green circle) thus inducing locus specific interactions (blue bars) and stabilizing the contacts between a specific enhancer, here CS4, and the *Igf2* gene. Altogether, at this locus, specific interactions alter the statistical shape of the chromatin fiber in chromosome- and tissue-specific ways, depending on the DNA methylation status of the ICR and on the TFs bound on the enhancers. Such chromatin allosteric effects allow the harmonious regulation of *Igf2* transcription (figure adapted from [61]).