Genes, Genomes and Genomics ©2007 Global Science Books

Transition from DNA Looping to Simple Binding or DNA Pairing in Gene Regulation and Replication: A Matter of Numbers for the Cell

Michèle Amouyal

CNRS-Interactions à Distance, 121 avenue Philippe Auguste – F75011 Paris, France

Correspondence: **michele.amouyal@club.fr**

ABSTRACT

Repression of the *E. coli* lactose operon is achieved though DNA looping and three operators at the physiological repressor concentration. In strains overproducing the repressor or with plasmids with a high copy number, the cooperative mode of repression is masked by other modes. *In vitro*, when several DNA molecules are present, DNA loop formation is replaced by intermolecular associations still mediated by the *lac* repressor. In bacteria, such associations, known as "handcuffing" and mediated by the initiator protein, are observed in replication of the iteron-class of plasmids. When moderate amounts of initiator and its binding to the replication origin (achieved in some instances, by DNA looping) allow replication to proceed, high concentrations prevent replication and lead to handcuffing that controls the number of plasmids. In principle, when DNA looping is feasible, DNA pairing is also possible if more than one DNA molecule is present in the cell. In eukaryotes, the action of the CTCF protein is particularly representative of this situation. This key component of elements that insulate gene expression from the surrounding genomic effects in vertebrates, also acts as an organizer of higher-order chromatin structures at the β-globin and *Igf2/H19* loci. At this latter locus, CTCF controls genomic imprinting by DNA looping. Recent data suggest that genomic imprinting and monoallelic expression might also be controlled through chromosome pairing.

Keywords: CTCF, lactose operon, oligomerization, plasmid replication

CONTENTS

Molecular biology has been first and normally interested in identifying the components that are involved in the basic cellular processes. When it was noted that some elements were in various concentrations, naturally or artificially, due to introduced genetic mutations, it was essentially descriptive. Thus, for DNA-protein interactions involved in gene expression and their regulation, the general idea was that the protein had to be in sufficient amount in the cell to ensure binding to the corresponding site and generate the process of interest. In bacteria for instance, it seemed unimportant to use strains overproducing a transcription factor instead of the wild-type strain and physiological amounts of protein. And because it is easier technically, multi-copy plasmids are often used in place of chromosomal constructs.

Since then, it has appeared in several instances that the number of copies of a gene or of a protein was able to change the regulation process.

One example is provided by the *E. coli* multi-site *lac* repression. From the approach employed to demonstrate the feasibility of DNA looping between *lac* repressor and two *lac* operators on the same molecule, it was clear that DNA looping could be concentration-dependent (Kramer *et al*. 1987; Amouyal 1991). When mediated by *lac* repressor, DNA looping is lost with increased concentrations of repressor and DNA molecules to the benefit of other interactions, independent binding to the operators or intermolecular associations. This finding was a decisive element towards the uncovering of the functionality of the secondary *lac* operator sites.

DNA looping is not restricted to transcription. It is also involved in replication, to initiate it for the bacterial R6K plasmid (Muckherjee *et al*. 1988; Miron *et al*. 1992) or on the contrary to stop it for the *E. coli* F-factor and to limit its number to one copy (cf. Zzaman and Bastia 2005). The initiator protein bridges these loops. However, the copy number of the iteron-class of plasmids is generally controlled by "handcuffing", i.e. intermolecular associations bridged by the initiator protein (Pal and Chattoraj 1988; McEachern *et al*. 1989).

Intermolecular associations are also represented by the pairing of chromosomes for the coordinated regulation of their genes. The CTCF protein is involved in genomic imprinting of the mammalian *Igf2/H19* locus. It does so by controlling the target of the enhancer shared by the two imprinted genes through alternate DNA looping and methylation of the CTCF binding site (Murrell *et al*. 2004). It also participates in chromosomal associations (Ling *et al*. 2006; Zhao *et al*. 2006).

The interplay between DNA looping, simple binding and DNA pairing is developed for these three specific situations.

Received: 28 February, 2007. Accepted: 16 April, 2007.

Invited Review

1. FROM THE NUMBER OF *LAC* OPERATOR SITES AND REPRESSOR MOLECULES TO DNA LOOPING AND TO MULTI-OPERATOR REPRESSION OF THE *E. COLI LAC* OPERON

It has been known since the origin of molecular biology (Jacob and Monod 1961) that the *lac* operon is repressed from a 21 bp operator sequence, *O1*, located on the promoter region and centered at +11 with respect to the transcription start. A protein, LacI, is responsible for this repression when it binds the operator (Gilbert *et al*. 1966, in a hypothesis formulated by the former ones). A simplified organization of the *lac* operon is shown in **Fig. 1**.

In the seventies, two homologous sequences, *O2* centered at +412 and *O3*, centered at -82, were discovered (Reznikoff *et al*. 1974; Gilbert *et al*. 1975).

During nearly 20 years, these sites have been regarded as cryptic sites, like the two promoter-like elements, P2 and P3, of the *lac* operon (Xiong *et al*. 1991). In fact, repression in the absence of the secondary operators is efficient, due to the strong affinity of repressor for *O1* ($K \sim 10^{13}\ M^{-1}$, according to Winter *et al*. 1981). Thus it was not clear why repression had to be assisted. Moreover, the constitutive mutations were all mapped in *O1* and none exclusively in the *O2* or *O3* regions.

This point was corroborated by the weak affinities of the repressor for *O2* and *O3*. *O2* has a ~10-fold reduced affinity compared to *O1* (Pfahl *et al*. 1979; Winter *et al*. 1981a) and *O3* an at least 100-fold reduced affinity (Pfahl *et al*. 1979; Winter *et al*. 1981a). A fragment carrying all three operators has nearly the same affinity for the repressor than a fragment carrying the sole *O1* operator (Pfahl *et al*. 1979; O'Gorman *et al*. 1980).

For these reasons, the possible contribution of *O2* or *O3* was discarded until the next decade, and a first decisive impulse with the discovery of the eukaryotic enhancers (Moreau *et al*. 1981; Banerji *et al*. 1981). Soon after, in 1983-1984, somewhat analogous sequences were found necessary for repression of two prokaryotic operons, the arabinose and galactose operons, in addition to a first operator sequence located in the vicinity of the promoter (Irani *et al*. 1983; Dunn *et al*. 1984).

As a result, two groups in 1984-1985 tried to establish a multi-site repression in an artificial *lac* systems of expression lacking the native *lac* operon (Herrin *et al*. 1984; Besse *et al*. 1986). These works summarize the elements that were supposed to be necessary for multi-site repression in the mid-eighties.

Since multi-site repression was observed in the *gal* operon with a control element within a structural gene reminding of *lacO2*, but not in the *lac* operon, the genomic and biochemical organization of the *gal* control region seemed of first importance. For this reason, the artificial regulation built by Herrin and Bennet 1984 mimicked the *gal* operon with a hybrid *lac-trp* promoter controlling *galactokinase* expression.

- In the native *lac* operon, *lacO1* overlaps with the promoter. In the *gal* operon, the *galOi* and *galOe* operators surround the promoter, allowing the RNA polymerase to sit on it in the presence of the repressor. Thus, a *trp* promoter was flanked by two *lacO1* operators. The *lacO1* operator was employed as a sequence with sufficient affinity to ensure operator occupancy, in a way similar to *galOe* and *galOi* when binding the *gal* repressor.
- The constructs were carried by a 20-copy plasmid.
- An effect was only observed when the repressor was provided by a $laci^{Q}$ strain, producing 10-fold more repressor than the native strain (Muller-Hill *et al*. 1968).

It was the first report that two *lac* operators can cooperate in repression. However, interpretation of the data is somewhat complicated by the interference with *trp* repression.

The artificial *lac* regulation built by Muller-Hill's group was inspired by Herrin and Bennet's work and also displayed some cooperativity. It is closer to a *lac* regulation since a *lac* promoter controlling *lacZ* expression replaced the *trp* promoter-operator on a plasmid. However, the *lacO2* and *lacO3* sites were not removed, since they were thought to be inactive. The *lac* promoter was surrounded by two "ideal" *lac* operator sites with a ten-fold increased affinity for the repressor, as compared to the *lacO1* site (Sadler *et al*. 1983). The study was carried out in a $laci^{Q1}$ strain producing even more repressor than the $laci^{Q}$ strain, about 100-fold more than the wt strain (Calos and Miller 1981).

Two "clichés" would have to be overcome to show the contribution of *O3* and *O2* to *lac* repression under native conditions. They are complementary because of their relation to the law of mass-action. First, ten copies of repressor are not sufficient to have the secondary sites occupied. These ten copies are the number of tetramers produced by the *wt* cell (Gilbert and Müller-Hill 1966). Thus, the concentration of repressor in the cell has to be increased to force the occupancy of the sites. Second, the affinity of *O2* for the repressor, without speaking of that of *O3*, is too weak to allow their occupancy. Thus, their affinity has to be increased to force the occupancy.

In any case, in 1984-1985, whether for the natural arabinose and galactose regulations, or for the artificial *lac* regulations (for lack of functional secondary operators), two operator sites either separated by 110-115bp (Irani *et al*. 1983; Herrin and Bennet 1984) or 220-240 bp (Dunn *et al*. 1984; Besse *et al*. 1986), were required to have full repression. How they cooperate for repression at these distances in *E. coli*, had to be explained.

DNA looping generated by the simultaneous binding of the repressor to the two distant sites was one possible model, but was regarded as most unlikely before 1986. The opposition to this model was strong, as R. Schleif recalls it* (Schleif 2003).

In fact, concerning LacI specifically, some early *in vitro* data rather seemed to favor other models, such as sliding of the repressor along DNA until the "physiological" *O1* site (Winter *et al*. 1981b). Other *in vitro* preliminary data in 1985 indicated that the repressor-operator complex was not retained by nitrocellulose in filter binding assays when 220 bp separated the two operators. A nucleosome-type structure where DNA is wrapped twice around LacI, could explain this unusual situation (Besse *et al*. 1986). Other nucleosome-type models were proposed later on, such as the "repressosome" that would form between the two *gal* operators separated by 114 bp, the *gal* repressor and various other proteins, the RNA polymerase, the CRP or HU proteins (see for example Kuhnke *et al*. 1989; Aki *et al*. 1996).

Thus DNA looping for *lac* repression specifically and for other regulations in general, was not obvious and its feasibility had to be demonstrated. The *lac* repressor was an ideal protein for that purpose because of its structure. This tetramer is organized into two dimers that bind DNA, as determined by various techniques (Kania and Brown 1976; O'Gorman *et al*. 1980; Culard and Maurizot 1981). It could then in principle induce DNA looping when mixed with a fragment carrying two *lac* operator sites.

In fact, according to the concentrations of repressor and DNA, various complexes, apart the loop, could be predicted for a mixture of repressor and fragment carrying two distant operator sites (Kramer *et al*. 1987; Amouyal 1991; **Fig. 2**).

Sufficiently low concentrations of DNA and repressor were expected to favor the simple binding (A) to one operator site. If flexibility of DNA allowed it, simultaneous binding to the second operator site and looping of the interven-

* "In 1984, I was regarding serenely the formation of DNA-protein loops because I wanted to prove the specific interaction between the CRP activator and RNA polymerase at the lac promoter (see conclusion of Amouyal and Buc 1987) by ligation experiments. This meant making loops between CRP and RNA polymerase. I transposed this project to the lac repressor in 1985, forgetting the ligation step for which it was not required. The initial project, as well as its transposition to the lac repressor, was widely spread before it was even started."

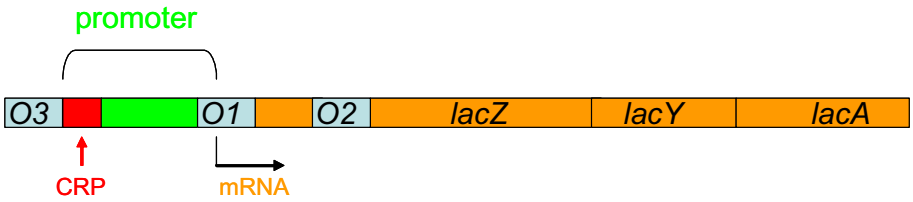

Fig. 1 Schematic map of the *E. coli* lactose operon (not to scale). Operators sites in blue, structural genes of the operon in yellow, promoter region in green, binding site for the CRP activator in red.

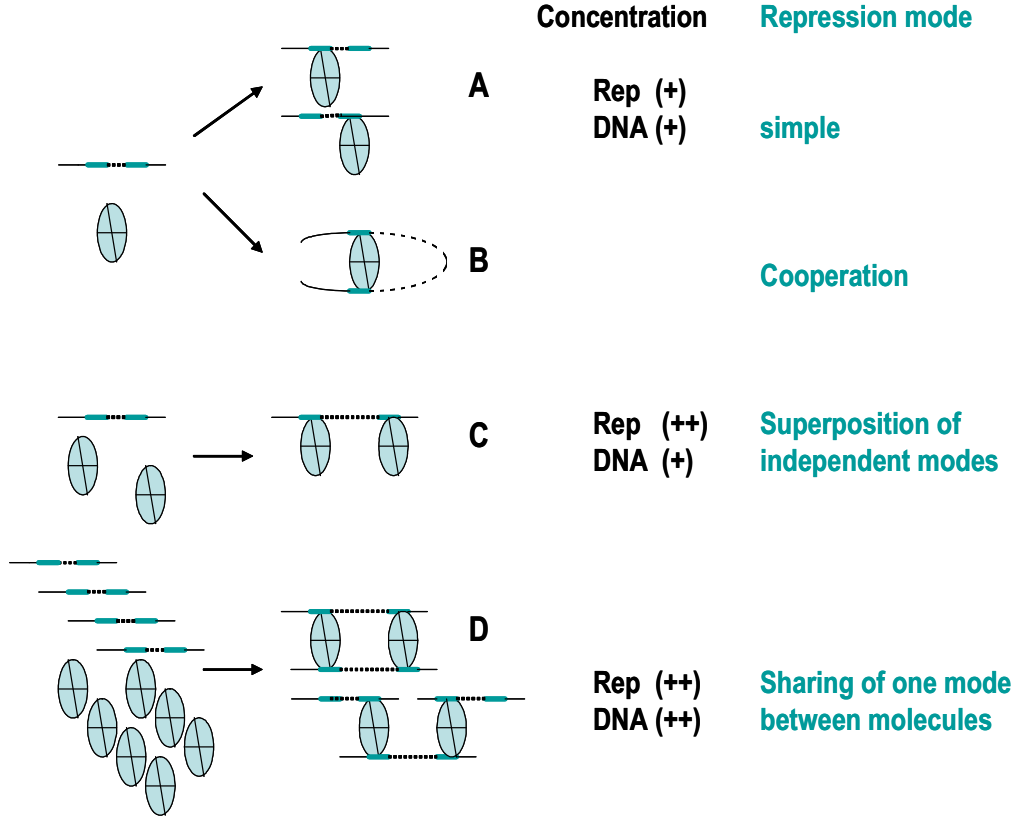

Fig. 2 The various repressor-operator interactions obtained between *lac* repressor and a fragment carrying two *lac* operator sites (first column) with increasing concentrations of repressor and DNA (second column). (**A**) Binding to only one operator site; (**B**) simultaneous binding of the repressor to the two operator sites with looping of the intervening DNA. (**C**) "tandem" binding to the two operator sites of the same fragment. (**D**) intermolecular associations. The third column indicates the expected repression mode.

ing DNA (B) would occur. For increased concentrations of repressor, the sites would be independently occupied in a "tandem" structure (C), and for both increased repressor and DNA concentrations, intermolecular species (D) where the repressor is in sandwich between two DNA molecules, could be expected.

The Kramer *et al*. (1987) study is the exact translation of this concentration scheme and of other anticipated loop features. The retardation gel assays† were performed so as the properties of the loops are directly read on the gels without any effort: continuity of DNA looping over a large range of distances, as opposed to punctual formation of nucleosome-type structures at 220 bp or other punctual structures, concentration effect on DNA loop formation, identification of intermolecular adducts, phasing at short distance as a consequence of DNA looping, if not a proof. Electron microscopy was reserved to the observation of large size loops and to the demonstration of an optimal distance for "DNA cyclisation" around 500 bp (the "cyclization" approach is detailed in Amouyal 1991).

The study clearly showed that DNA looping is favored *in vitro* by the concentrations that do not saturate the operators with repressor. In the w.t. strain, there is only one copy of DNA per cell, the chromosome, and the cell produces only 10 copies of tetrameric repressor (Gilbert and Muller-Hill 1966). The concentration of repressor can be increased 10-fold in a $laci^{Q}$ strain (Calos and Miller 1981) and about 100-fold in a $laci^{Q1}$ strain (Müller-Hill *et al*. 1968). The number of operator copies was 20 in the Herrin and Bennet (1984) study and can reach 500-700 copies per cell for the pUC-based plasmids. Thus, the use of strains overproducing the repressor and of multi-copy plasmids might have concealed DNA looping in the former artificial *lac* regulation. To detect DNA looping, if it did occur, the assays had to be carried out as close as possible to the native conditions, with the three operators on the chromosome (or very low copy plasmids, at most) and with the repressor produced by the chromosomal *laci* gene.

The second "cliché" was related to the supposedly too low affinity of the secondary sites for them to be implied in repression. It was relying on a fact known for long and confirmed by several groups: the association of the repressor with a fragment carrying all three operators is hardly stronger (four-fold or less) than with a fragment carrying the single *O1* operator (Pfahl *et al*. 1979; O'Gorman *et al*. 1980). DNA supercoiling broke this "cliché". My work of transition between Chemistry and DNA looping (Amouyal and Buc 1987) had made me aware that DNA is supercoiled in *E. coli*. If DNA looping was feasible on a linear template, DNA supercoiling could perturb it, or on the contrary, assist it. Thus, from the beginning of my work on DNA looping in 1984-1985, a second step was planned with a supercoiled template (cf. Kramer *et al* 1988). In fact, *in vitro*, DNA looping is strongly stabilized by DNA supercoiling: the half-life of the *O1-O2-O3* complex on a plasmid is 28 h instead of 7 min on a linear template (Whitson *et al*. 1987). On a plasmid, the secondary operators slow LacI dissociation 200-fold according to Whitson *et al*. (1987), from 7-fold to 32-fold according to Eismann and Müller-Hill (1990), when comparing the half-lives of the *O1-O2-O3* and *O1* complexes. Direct *in vivo* approaches have also unravelled the influence of DNA supercoiling. *In vivo* DNA footprinting of the *O1-O3* region by various chemical agents shows that *O3* is only occupied when DNA is supercoiled (Boro-

† The preliminary data with the retardation gel experiments were communicated to the Berlin Summer School on DNA structure in 1986.

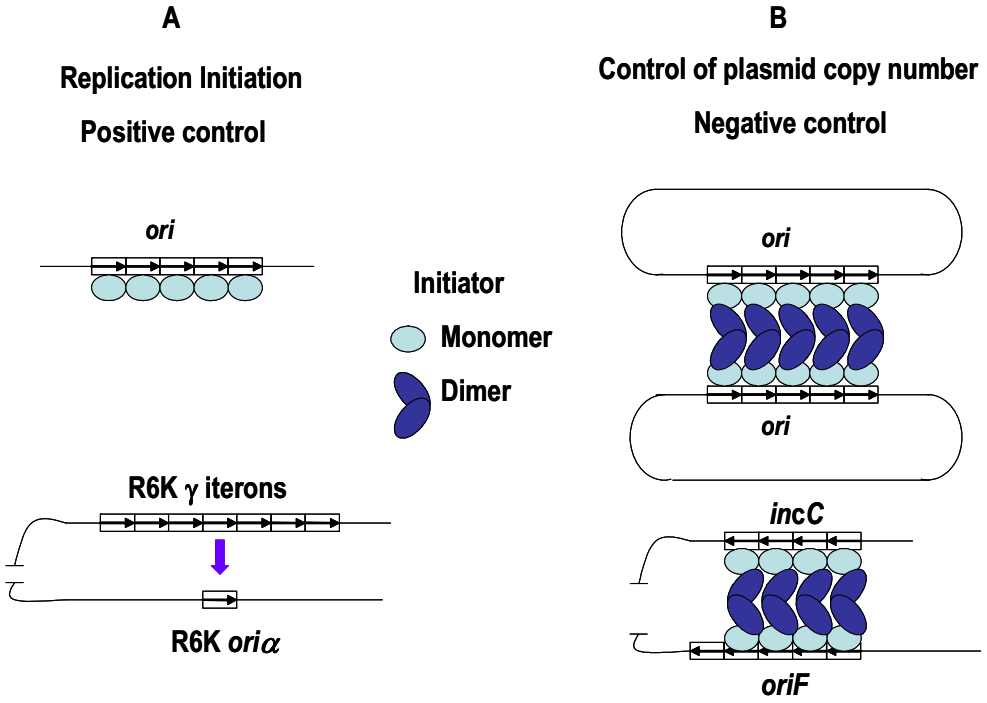

Fig. 3 Replication of the iteron-class of plasmids. Panel A: cooperative binding of the initiator to the iterons for replication initiation (general case). Below, particular situation of the R6K plasmid with DNA looping for initiation of R6K at the α (or β) origin. **Panel B**: "handcuffing" for the control of the plasmid copy number (general case). Below, particular situation of the F-factor with DNA looping for limitation of its copy number to one.

wiec *et al*. 1987, see also Flashner and Gralla 1988 for the *in vivo O1-O2* occupancy). From these assays, Gralla and co-workers have deduced that *O3* binds the repressor *in vivo* only 10-fold less tightly than *O1*, whereby *in vitro* independently on a linear template, O3 binds the repressor at least 100-fold less tightly (Pfahl *et al*. 1979; Winter *et al*. 1981a). DNA supercoiling makes DNA loop formation so easy that the requirement for relative orientation of two *lac* operators that is obvious between 153 and 168 bp on a linear or relaxed DNA template, is no longer legible on a supercoiled one. Instead, loops are formed for all distances and only their stability, instead of their existence, is now sensitive to helical orientation of the sites (Kramer *et al*. 1988).

In spite of these results, the clichés took some time to disappear since all the *lac* regulations performed before 1990 still included the O3 site as a cryptic site thought inactive, or were run with strains overproducing the repressor and plasmids with a high level of copies, which somewhat perturbs the interpretation of their data.

Finally, the inactivation of each site by mutagenesis, including the *O3* site under conditions as close as possible to the w.t. ones, with the chromosomal repressor gene and the three sites on the chromosome, determined the contribution of each operator site to repression (Oehler *et al*. 1990, 1994). The picture of *lac* repression was definitely modified. It appeared that repression of the *lac* operon relied on the three operators and not exclusively on *O1*. This contribution depends on the intracellular concentration of *lac* repressor. When the three operators are on the chromosome, repression enhancement at *O1* by both *O2* and *O3*, decreases from 72-fold to 48-fold and 6-fold when the number of *lac* repressor tetramers increases from 10 to 50 and 900, respectively. In fact, repression issued from *O1-O2* or *O1-O3* cooperation is replaced by repression from independent binding at *O1*, *O2* and *O3*.

This view is supported by the results obtained with the single operators and also by the use of repressor mutants that are dimers unable to form tetramers and to induce DNA looping (Alberti *et al*. 1991; Chakerian *et al*. 1991). Thus at relatively low repressor concentration (200 subunits of either dimer or tetramer), the dimer induces the same level of repression from the three operators than the tetramer from the single *O1* operator (Oehler *et al*. 1990; Oehler *et al*. 1994). This is the indication that, when the repressor is unable to form DNA loops, it hardly binds *O2* and *O3* at low, near-to-physiological intracellular repressor concentrations.

This is an over-all effect resulting from alternative looping between either *O1* and *O2,* or *O1* and *O3*. This concentration effect is more clearly followed when only one loop is present, either the *O1-O2* loop or the *O1-O3* loop and when only two operators are present and the third one is inactivated.

In a situation where only the *O2* and *O1* operators are present on the chromosome, the repression enhancement of *O2* versus *O1*, falls from a 10/11.5-fold effect to none (1.3/1.5 fold effect) with an increasing amount of tetramers from 50 to 900 molecules (Oehler *et al*. 1990, 1994). The influence of *O2* is also lost with elevated repressor concentrations and when *O2* or *O3* replace *O1* on the promoter (Oehler *et al*. 1994). On the contrary, when the dimer replaces the natural tetrameric repressor, i.e. in the absence of DNA looping, there is no influence of *O2* at 400 bp of a first operator overlapping the promoter, and this situation is not concentration-dependent.

The *O1-O3* interaction is less simple than the *O1-O2* interaction. Indeed, the protein responsible for activation of the *lac* operon, the CRP protein, binds a site centered at -61.5 close to the *O3* site (-82). Is CRP responsible for the repression observed from *O3*, whether by competing with repressor for binding at the same location or by assisting repressor binding (Fried and Hudson 1996; Perros and Steitz 1996)?

In the *lacZ* constructs of Amouyal and von Wilcken-Bergmann (1992), there is no site for CRP binding and artificial *lacWM* sites have been used either on the promoter, at the *O1* location, or upstream, at the *O3* location. The *lacWM* sites do not allow repression by themselves when located either on the promoter or upstream. However, when two *lacWM* operator sites are in place of *O1* and *O3*, the β-galactosidase activity is repressed from 8- to 35-fold, depending on their relative orientation and on whether the repressor gene and operator sites are carried by the chromosome or 5-copy plasmids. This significant level of cooperation was reproduced with two *O3* operators (Oehler *et al*. 1994). Furthermore, the cooperation resulting from the exchange of *O1* by *O3* in the originally *O1-O2* construct (Oehler *et al*. 1994) is again an indication that DNA looping involves *O3* without the need for CRP. Thus, unassisted DNA looping is a true component of the naturally observed repression between *O1* and *O3*. The concentration effect is also consistent with DNA looping since the effect of *O3* with respect to *O1* falls from 28-fold to 5.6-fold from 200 to 3600 subunits

of tetramer. With the dimer, this effect remains in the lower range. Then, DNA looping competes with the activation process, leading to an additional level of repression (Oehler *et al*. 1994).

Note that in the artificial regulation described in Amouyal and von Wilcken-Bergmann (1992) or Perez *et al*. (2000), there is no need for assistance, except for natural supercoiling, whether two "poor" *lac* operator sites cooperate (Amouyal and von Wilcken-Bergmann 1992) or two proteins with good affinity for their sites, interact weakly, like the two galactose repressor dimers (Perez *et al*. 2000).

According to a recent work by Zhang *et al*. (2006), DNA looping between two *lac* operators separated by 100 bp would require the nucleoid HU protein. It should be noted that interpretation of the data obtained with HU-deficient strains is uneasy because of various mutations generated by the bacterium to compensate for this deficiency, as pointed out in Perez *et al*. (2000) and in Amouyal (2005). Second, DNA supercoiling is sufficient to explain DNA looping with two *lac* operators as above detailed (see also Purohit and Nelson 2006).

2. REPLICATION OF THE ITERON-CLASS OF PLASMIDS AND CONTROL OF THEIR COPY NUMBER IN BACTERIA

At an international conference where I was presenting the first data on DNA looping with the *lac* repressor, specialists of plasmid replication came to me and made the remark that the concentration scheme displayed in **Fig. 3** could as well be applied to the control of plasmid copy number (cf. Pal and Chattoraj 1988; Nordstrom 1990).

In fact, the number of plasmids, these extra-chromosomal genetic elements in bacteria, is kept within precise limits in the cell. There are for example 1 to 2 P1 plasmids per cell (cf. Das *et al*. 2005), about 10 R6K plasmids (cf. Muckergee *et al*. 1988), about 15 colE1 plasmids (cf. Twigg and Sherratt 1980) and only one F-factor (cf. Zzaman and Bastia 2005) in *E. coli*, 15 to 18 pPS10 plasmids in *Pseudomonas aeruginosa* (cf. Giraldo and Fernandez-Tresguerres 2004).

The use of multiple short repeats of DNA (of about 20 pb), or iterons, is a major way of controlling the number of plasmids in Gram-negative bacteria. This family includes the R6K, P1, F and pSC10 plasmids, among others. These iterons are present at the replication origin. Their replication presents some analogy with the transcriptional repressors.

To control the number of these plasmids once a specific number of copies has been reached, the bacterium prevents their replication.

Replication is inhibited by plasmid pairing at the origin of replication, called "handcuffing" (Pal and Chattoraj 1988; McEachern *et al*. 1989; **Fig. 3B**) as detected by ligation experiments (Kunnimallaiyaan *et al*. 2005), the actual basis for the 3C technique widely used in eukaryotic cells to detect DNA looping or pairing (see section 3).

DNA pairing is bridged by a key component of replication, the initiator protein. The initiator is indeed first required for initiation of replication at the iterons, by cooperative binding to the iterons (**Fig. 3A**). Handcuffing prevents new rounds of plasmid replication and is linked to the increasing concentration of initiator with the number of plasmids. Thus, the number of Pi dimers, the initiator of R6K replication, can reach about 6000 molecules per cell (Filutowicz *et al*. 1986).

For some plasmids of the iteron family, such as the R6K and P1 plasmids, there is another set of DNA repeats, at some distance from the origin of replication, in a region that causes incompatibility, i.e. the impossibility for two groups of plasmids to co-exist in the same cell. This distant region is named *inc* for P1, and γ for R6K. In fact, the bacterium applies the same procedure, handcuffing, when another plasmid carrying the same replication elements is introduced and increases the global number of plasmids in the cell.

In case of the *E. coli* F-factor, handcuffing is replaced by DNA looping between the origin of replication and *incC*, at about 1.5 kb (**Fig. 3B**). DNA looping inhibits strand separation required for initiation of replication at the origin (Zzaman and Bastia 2005). This allows the cell to limit the number of copies to a single one in case of the F-factor.

Thus, the initiator is both involved in the positive and the negative control of replication, like the transcriptional activator and repressor protein, AraC (Schleif 2003). And in a similar way, conformational transitions modify the function of the protein. Additionally to a repertoire of mutations for all known initiators, three structures have been solved: the iteron-RepE monomer complex from F-factor (Komori *et al*. 1999), the dimerization domain of pPS10 RepA (Giraldo and Fernandez-Tresguerres 2004) and the iteron-Pi monomer complex (Swan *et al*. 2006). This provides a wealth of information and a global view of how the initiator acts.

In the solution, the initiator exists predominantly under the dimeric form deprived of initiation activity. In this form, it binds poorly the iterons and cannot activate replication from the iterons. It can, however, bind an inverted repeat, in which the two half-sites of the iteron are in opposite directions. When this inverted repeat is present, adjacent to the iterons of the origin of replication, the initiator is also able to repress its own synthesis (see for example replication control region of the pPS10 plasmid; Giraldo and Fernandez-Tresguerres 2004).

A chaperone system generally mediates the conversion of the inactive dimeric form into an active one for initiation (Zzaman *et al*. 2004). The active form is a monomer actually oligomerized by its binding to the iterons (Germino and Bastia 1983).

The two forms of the protein, monomeric and dimeric, are involved in DNA looping or handcuffing. Contrary to the transcriptional repressors, the monomer has the peculiarity to bind the two half-sites of the iteron both through the N-terminal and the C-terminal domains, as observed for the F RepE-iteron (Komori *et al*. 1999) and R6K Pi-iteron crystals (Swan *et al*. 2006).

For some plasmids, such as the R6K family, this is not the end of the story and replication provides an example of the gymnastics that the same molecular elements can perform to achieve different goals in the cell with an economy of means (cf. Schleif 2003). Indeed, for the R6K plasmids carrying the distant γ set of iterons, replication from the sole origin of replication is not efficient. Replication requires the distant γ iterons to proceed. From this point of view, the γ iterons are also replication enhancers.

More precisely, the replication proceeds from two origins, α and β. The γ element contains an array of seven 22 bp repeats that can bind the initiator protein. The α origin contains a binding site for the initiator, but it binds the protein poorly by itself. Initiation at this origin can only start when the seven γ iterons, about 4 kb away (the largest distance observed for enhancer action in *E. coli*), are physically linked to the α site through the initiator (**Fig. 3A**). This contact stabilizes the protein at the α origin. In the same way, the β origin contains a half-iteron only involved in initiation when at the same time, the initiator occupies the γ iterons, about 1.2 kb away (Muckherjee *et al*. 1988; Miron *et al*. 1992; Swan *et al*. 2006).

3. CONTROL OF GENOMIC IMPRINTING BY DNA LOOPING OR CHROMOSOME PAIRING IN MAMMALS

From the above examples related to bacteria, it is clear that when DNA looping is possible *in cis*, mediated by a protein, *trans*-associations are also possible with the same molecular elements when several DNA molecules are present, in order to achieve the same biological process. Recent findings related to genomic imprinting of the mammalian *Igf2/H19* locus and CTCF, indicate that this is not restricted to *E. coli*.

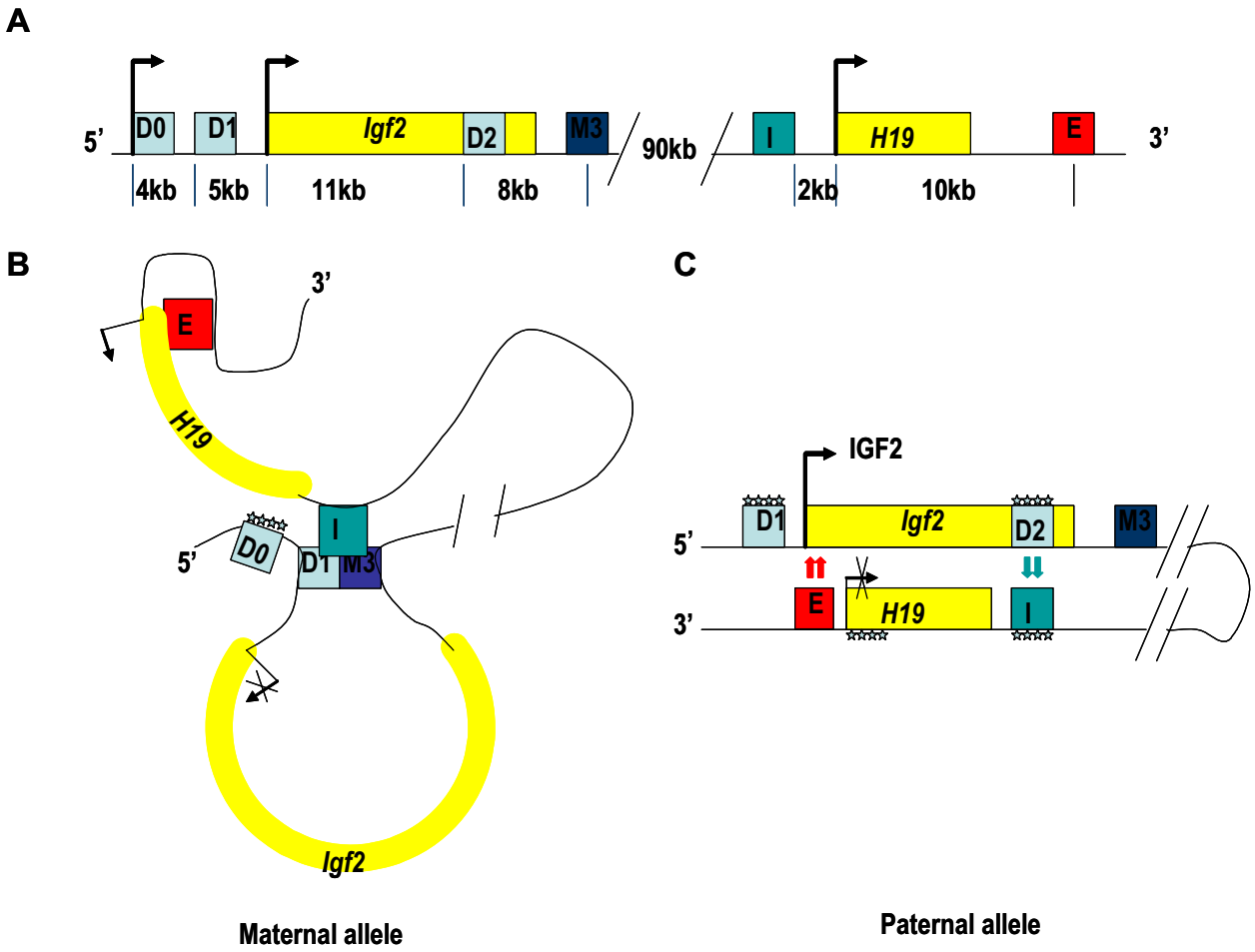

Fig. 4 Panel A: Simplified view of the mouse *Igf2/H19* locus. The arrows indicate the promoters of the two genes, *Igf2* on the left, *H19* on the right, in yellow. The enhancers shared by the two genes are indicated by a red E-square, the imprinted control region (ICR), in medium blue, the differentially methylated domains, DMR0, DMR1and DMR2, in light, the matrix attachment region, MAR3, in dark blue. **Panel B**: the spatial architecture induced by CTCF (and possibly other proteins) when it binds to its unmethylated sites on the ICR region and on the DMR1 region, on the maternal allele. As a result, the enhancer can only contact and switch on the H19 gene whereas the *Igf2* gene, trapped in one of the loops, is repressed. **Panel C**: On the paternal allele, the CTCF sites are methylated. The asterisks underline the methylation status of DNA. The ICR-DMR1-MAR3 interaction of the maternal allele, is replaced by an ICR- DMR2 interaction, bringing the enhancer closer to the *Igf2* promoter and switching on the *Igf2* gene, whereas DNA methylation represses the *H19* promoter.

The CTCF (CCCTC-binding factor) protein was first found as part of the 5'-HS4 insulator element of the chicken β-globin locus, separating the active globin domain from the heterochromatic domain. The 5'-HS4 element protects gene expression against enhancers when placed between the gene and enhancer, as well as against position effects on the chromosome (see for example Burgess-Beusse *et al*. 2002). CTCF is required for the enhancer blocking activity (Bell *et al*. 1999).

This activity is a general feature of the CTCF protein that is highly conserved in higher eukaryotes, from *Drosophila* to humans (Moone *et al*. 2005).

The 11 zinc fingers of the protein allow its binding to a multiplicity of DNA targets (cf. Ohlsson *et al*. 2001). CTCF binds repeated sites in a highly cooperative manner (Pant *et al*. 2004). In the absence of DNA, it can also dimerize with help of the zinc fingers or the C-terminal part of the protein (Pant *et al*. 2004).

In mouse erythroid cells, the CTCF-binding sites participate in spatial interactions between the active β-globin genes and, about 50 kb away, the LCR required for their high level expression. This results in what was termed by the authors an Active Chromatin Hub with several active loops acting in concert over 200 kb (Tolhuis *et al*. 2002). The CTCF binding sites seem to structure an inactive chromatin pre-framework, as was found from analysis of the β-globin locus in progenitor cells, where the globin genes are not yet active (Splinter *et al*. 2006).

CTCF also seems to organize spatially the coordinated regulation of different genes and loops at the mammalian *Igf2/H19* locus, and to control their genomic imprinting in this way (see for example, Reik and Walter 2001).

Thus, the insulin growth factor, IGF2, is a regulator of fetal growth. The corresponding gene is imprinted, i.e. it is only expressed from one of the two parental alleles. Its expression is coordinated with that of another imprinted gene, *H19*. The two genes are located 90 kb apart on mouse chromosome 7 (**Fig. 4A**). On the maternal chromosome, *H19* is expressed whereas *Igf2* is repressed throughout development. The opposite is found on the paternal allele. In mouse neonatal liver, transcription of the two genes is regulated by an enhancer located downstream of *H19* (**Fig. 4A**), about 10 kb from *H19* and about 120 kb from *Igf2*. A cluster of four binding sites for CTCF is located between the two genes, in the ICR imprinting control region that regulates the methylation of different sites within the cluster and is located 2 kb upstream of *H19*. CTCF cannot bind to its sites when they are methylated.

On the maternal inherited chromosome, the differentially methylated region, DMR1, also contains binding sites for CTCF. It physically interacts with the unmethylated ICR though CTCF and possibly other proteins (Kurukuti *et al*. 2006), as detected by the 3C technique (Dekker *et al*. 2002). The ICR also physically interacts with the matrix attachment region MAR3 located next to the DMR1 domain. *Igf2* is trapped in a 20 kb loop (**Fig. 4B**). This is seemingly sufficient to prevent the enhancer to activate the maternal *Igf2* promoter (Kurukuti *et al*. 2006).

Previously, the same group had shown that on the paternal chromosome, the methylated ICR physically interacts with one of the differentially methylated region, DMR2, of *Igf2* (Murrell *et al*. 2004). This brings the *Igf2* promoter into contact with the enhancer in a 110 kb DNA loop, thus leading to *Igf2* expression (**Fig. 4C**). As the enhancer is occupied elsewhere, it cannot activate *H19*. The proteins responsible for the ICR-DMR2 interaction have not been identified yet. Some zinc-finger proteins bind methylated DNA. Some of them colocalize with the *H19/Igf2* ICR (Fillion *et al*. 2006). These proteins have been found to assist repression of *H19*. In analogy with CTCF, they might also bridge the ICR-DMR2 interaction.

Thus CTCF-mediated DNA looping and methylation jointly direct the enhancer to a different promoter target on the paternal and maternal chromosomes.

Since CTCF mediates DNA looping, it is not surprising, at least with the viewpoint of this review, that CTCF is also involved in intermolecular associations.

Accordingly, Ling *et al*. (2006) have recently found that the *Igf2/H19* locus on mouse chromosome 7 was physically linked to another locus, *Wsb1/Nf1*, on chromosome 11 via CTCF bound to the Imprinting Control Region. This association was found by screening for all CTCF-mediated associations by the 3C technique (or a variant) and by fluorescent *in situ* hybridization (FISH). This is an intriguing fact because only the paternal allele of chromosome 11 is found though this locus is not imprinted. Deletion of the maternal ICR cancelled the association. The silencing of CTCF by RNA interference disrupted the coordinate regulation of the two loci. Only expression of the paternal *Wsb1/Nf1*locus is altered. Thus it was concluded that the association and CTCF allow *trans*-regulation of the genes or alternatively that the two chromosomes share separately the same elements concentrated at a transcription factory (Hugues *et al*. 1995; Chakalova *et al*. 2005).

Since the association is restricted to the paternal allele of chromosome11, it might be involved in the imprinting process of the *Igf2/H19* locus on chromosome 7 (Ling *et al*. 2006), or in a transient methylation and silencing of the *Wsb1* gene at the *Wsb1/Nf1* locus (Krueger and Osborne 2006).

Similar data have been qualitatively obtained by Zhao *et al*. (2006) with a 3C-based technique, except that they find 114 inter-or intra-chromosomal associations instead of 3 for Ling *et al*. (2006). Up to four chromosomes converge to the ICR of the *Igf2/H19* locus. Clearly, their experimental conditions increase the sensitivity of the technique. It is also clear that the uncovered interactions are preferentially linked to imprinting. They are also specific of a certain stage of growth and depend on the reprogramming of the cells.

Inter-chromosomal associations might present some similarity with transvection, a phenomenon known since 1954 in *Drosophila*. Transvection is an alteration of gene expression that depends upon whether or not genes are paired with their homologs (Wu 1993; Duncan 2002). It generally involves the action of enhancers in *trans*.

CTCF binding sites sensitive to DNA methylation have also been found near the 3′ end of *Xist*, the gene on the X-chromosome that triggers the inactivation of a single X-chromosome in X-X female mammals (Chao *et al*. 2002), as well as at the boundaries of domains which escape X-inactivation (Filipova *et al*. 2005). The precise role of CTCF in both cases is not known. However, it recently appeared from two recent studies (Bacher *et al*. 2006; Xu *et al*. 2006) that the two X-chromosomes are transiently paired, as assessed by RNA-FISH analysis. This pairing precedes X-chromosome inactivation. Binding competition assays with sub-fragments of the X inactivation center region, disrupts X-inactivation and confirms this association. Competition with DNA fragments containing the CTCF binding site also disrupts X-inactivation (Donohoe *et al*. 2007). Like for the *Igf2/H19* ICR associations, this suggests a role for CTCF as a bridging protein, along with cofactors such as the Yy1 protein (Donohoe *et al*. 2007).

Another *trans*-chromosomal regulation was previously described by Spilianakis *et al*. (2005). This group has found that the interferon-γ promoter region on mouse chromosome 10 was juxtaposed with the locus control region of the interleukine genes on chromosome 11 in thymocytes precursors of the immune system T-cells, both CD4-helper T-cells and CD8-killer T-cells. It now appears that this inter-chromosomal association is lost when naive T-cells differentiate into T_H1 or T_H2 cells. On the contrary, in activated T_H2 cells, the production of interleukin is associated with the folding of the interleukin gene into several loops with coordinated expression, due to a protein, SATB1, only found in thymocytes. A cage-like distribution of SATB1 surrounds the nucleus and anchors the loops to the chromosomal scaffold (Cai *et al*. 2006). As the naive T-cells either do not express the cytokines or from only one allele under specific conditions, Spilianakis *et al*. (2005) suggest that pairing is associated with silencing of the allele.

Thus in mammals, chromosomal associations seem to promote silencing and monoallelic expression, whereby DNA looping is more ambivalent. This is also what emerges from bacterial replication, where handcuffing is exclusively devoted to silencing and to the control of plasmid copy number, whereby DNA looping can be used for both positive (cf R6K) and negative control (cf. F-factor) of replication, like in transcriptional regulation.

4. OLIGOMERIZATION OF THE PROTEIN

Several proteins involved in DNA looping are naturally oligomerized, sometimes to a high degree.

Thanks to a leucine zipper, two subunits of *lac* repressor interacting side-by-side as a dimer, can also interact head to head as a tetramer with two binding sites for DNA (Alberti *et al*. 1991; Chakaherian *et al*. 1991).

The *deo* repressor is a triple dimer, with three binding sites for DNA (Mortensen *et al*. 1989), resulting in double loop formation on binding with native DNA (Amouyal *et al*. 1989).

The λCI repressor is a dimer capable of tetramerization and octamerization in the absence of DNA (Senear *et al*. 1993; Bell *et al*. 2000; Bell and Lewis 2001). The octamerization allows DNA looping between the O_L and O_R arrays of repeats (Dodd *et al*. 2001).

The record of oligomerization pertains to the 186CI repressor that comprises 14 subunits as a heptamer of dimers arranged like a wheel with seven possible binding sites for DNA (Pinkett *et al*. 2006). Electrostatic forces rather than specific sequence determinants are responsible for this oligomerization. This oligomerization explains well the all-or-none occupancy of the three 186CI repeats of the phage P_R promoter (Dodd *et al*. 1996; Pinkett *et al*. 2006). The arrays of sites are wrapped around this "wheel", making the subunit organization particularly suited to the tandem arrangement of the sites (cf. Amouyal *et al*. 1998). For the 186CI repressor (and to a lesser extent, for the λCI repressor), it seems that the cell has selected both site and protein organization, in a way they are best fitted for adjacent binding.

Now, apart from assisting repressor binding to DNA repeats, oligomerization of the protein, when it provides an already assembled protein bridge, can also stabilize the loop, as is the case for the *lac* repressor. However, more than with a long range distance action, oligomerization in the absence of DNA seems to be associated with the facilitated recognition of multiple loci scattered on a precise portion of the genome, as appears from the *lac*, *deo* or *186CI* regulations. Indeed, if the two arrays of O_L and O_R sites of the λphage, are 2.3 kb distant, the most distant sites for the 186CI protein, F_L and F_R, are "only" 300 bp distant, and between these two sites, there are at least four other operator sites, close to P_L and P_R, 62 bp distant. The three *lac* operators are likewise scattered over 500 bp and the three *deo* operators over 880 bp.

In *E. coli*, the largest distance involved in DNA looping

is that observed for replication initiation of plasmid R6K. Specifically, the α origin iteron and the seven γ iterons bridged by Pi, are separated by 4 kb (**Fig. 3A**). With the present state of knowledge, it is difficult to specify the precise structure of the protein bridge responsible for this long distance action. However, DNA looping might require a specific initiator interface, different from the contacts that are already known between two initiator molecules (Swan *et al*. 2006).

Handcuffing seems to require repeated iterons, from 3 (pSC101) to 7 (R6K-γ) repeats, as well as strong interactions that are provided by a highly cooperative organization of the protein at the iterons with a layer of dimeric protein holding apart two layers of DNA-bound monomers (Toukdarian and Helinski 1998; Zzaman and Bastia 2005; **Fig. 3B**).

REFERENCES

Alberti S, Oehler S, von Wilcken-Bergmann B, Kramer H, Muller-Hill B (1991) Dimer-to-tetramer assembly of Lac repressor involves a leucine heptad repeat. *New Biologist* **3**, 57-62

Aki T, Choy HE, Adhya S (1996) Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of *gal* transcription by GAL repressor. *Genes Cell* **1**, 179-188

Amouyal M, Buc H (1987) Topological unwinding of strong and weak promoters by RNA polymerase. *Journal of Molecular Biology* **195**, 795-808

Amouyal M (1991) The remote control of transcription, DNA looping and DNA compaction. *Biochimie* **73**, 1261-1268

Amouyal M (2005) Gene regulation at-a-distance in *E. coli*: new insights. *Comptes Rendus Biologie* **328**, 1-9

Amouyal M, Wilcken-Bergmann B (1992) Repression of the *E. coli lactose* operon by cooperation between two individually unproductive 'half-operator' sites. *Comptes Rendus de l'Académie des Sciences* **315III**, 403-407

Amouyal M, Perez N, Rolland S (1998) Adjacent cooperation of proteins on DNA is not representative of long-distance interactions. *Comptes Rendus de l'Académie des Sciences* **321III**, 877-881

Amouyal M, Mortensen L, Buc H, Hammer K (1989) Single and double loop formation when deoR repressor binds to its natural operator sites. *Cell* **58**, 545-551

Bacher CP, Guggiari M, Brors B, Augui S, Clerc P, Avner P, Eils R, Heard E (2006) Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nature Cell Biology* **8**, 293-299

Banerji J, Rusconi S, Schaffner W (1981) Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299-308

Bell AC, West AG, Felsenfeld G (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387-396

Bell CE, Frescura P, Hochschild A, Lewis M (2000) Crystal structure of the λ repressor C-terminal domain provides a model for cooperative operator binding. *Cell* **101**, 801-811

Bell CE, Lewis M (2001) Crystal structure of the λ repressor C-terminal domain octamer. *Journal of Molecular Biology* **314**, 1127-1136

Besse M, van Wilcken-Bergmann B, Müller-Hill B (1986) Synthetic *lac* operator mediates repression through *lac* repressor when introduced upstream and downstream from *lac* promoter. *The EMBO Journal* **5**, 1377-1381

Borowiec JA, Zhang L, Sasse-Dwight S, Gralla JD (1987) DNA supercoiling promotes formation of a bent repression loop in *lac* DNA. *Journal of Molecular Biology* **196**, 101-111

Burgess-Beusse B, Farrell C, Litt M, Mutskov V, Recillas-Targas T, Simpson M, West A, Felsenfeld G (2002) The insulation of genes from external enhancers and silencing chromatin. *Proceedings of the National Academy of Sciences USA* **99 (Suppl. 4)**, 16433-16437

Cai S, Lee CC, Kohwi-Shigematsu T (2006) SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokines genes. *Nature in Genetics* **38**, 1278-1288

Calos MP, Miller JH (1981) The DNA sequence change resulting from the iq1 mutation which greatly enhances promoter strength. *Molecular and General Genetics* **183**, 559-560

Chakerian AE, Tesmer VM, Manly SP, Brackett JK, Lynch MJ, Hoh JT, Matthews KS (1991) Evidence for leucine zipper motif in lactose repressor protein. *The Journal of Biological Chemistry* **266**, 1371-1374

Chakalova L, Debrand E, Mitchell JA, Osborne CS, Fraser P (2005) Replication and transcription: shaping the landscape of the genome. *Nature Reviews Genetics* **6**, 669-677

ChaoW, Huynh KD, Spencer RJ, Davidow LS, Lee JT (2002) CTCF, a candidate *trans*-acting factor for X-inactivation choice. *Science* **295**, 345-347

Culard F, Maurizot JC (1981) Lac Repressor-*lac* operator interaction: circular dichroism study. *Nucleic acids Research* **9**, 5175-5184

Das N, Valjanec-Gratian M, Basuray AN, Fekete RA, Papp PP, Paulsson J, Chattoraj DK (2005) Multiple homeostatic mechanisms in the control of P1 plasmid replication. *Proceedings of the National Academy of Sciences USA* **102**, 2856-2861

Dekker J, Rippe K, Dekker KM, Kleckner N (2002) Capturing chromosome conformation. *Science* **295**, 1306-1311

Dodd IB, Perkins AJ, Tsuemitsidis D, Egan JB (2001) Octamerization of λCI repressor is needed for effective repression of PRM and efficient switching from lysogeny. *Genes and Development* **15**, 3013-3022

Dodd IB, Egan JB (1996) DNA binding by the coliphage 186 repressor binding protein CI. *Journal of Biological Chemistry* **271**, 11532-11540

Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT (2007) Identification of a CTCF cofactor Yy1 for the X chromosome binary switch. *Molecular Cell* **25**, 43-56

Duncan IW (2002) Transvection effects in Drosophila. *Annual Review of Genetics* **36**, 521-56

Dunn TM, Hahn S, Ogden S, Shleif RF (1984) An operator at -280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proceedings of the National Academy of Sciences USA* **81**, 5017-5020

Eismann ER and Muller-Hill B (1990) *lac* repressor forms stable loops *in vitro* with supercoiled at wt *lac* DNA containing all three natural operators. *Journal of Molecular Biology* **213**, 763-775

Filipova GN, Cheng MK, Moore JM, Truong JP, Hu YJ, Nguyen DK, Tsuchiya KD, Disteche CM (2005) Boundaries between chromosomal domains of X-inactivation and escape bind CTCF and lack CpG methylation during early development. *Developmental Cell* **8**, 31-42

Fillion G, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez PA (2006) A family of human zinc finger that bind methylated DNA and repress transcription. *Molecular and Cellular Biology* **26**, 169-181

Filutowicz M, Mc Easchern MJ, Helinski DR (1986) Positive and negative role of an initiator protein at an origin of replication. *Proceedings of the National Academy of Sciences USA* **83**, 9645-9649

Flashner Y, Gralla JD (1988) Dual mechanism of repression at a distance in the *lac* operon *Proceedings of the National Academy of Sciences USA* **85**, 8968-8972

Fried M and Hudson JM (1996) DNA looping and *lac* repressor-CAP interaction. *Science* **274**, 1930-1931

Germino J, Bastia D (1983) Interaction of the plasmid R6K-encoded replication initiator protein with its binding sites on DNA. *Cell* **34**, 125-134

Gilbert W, Muller-Hill B (1966) Isolation of the *lac* repressor. *Proceedings of the National Academy of Sciences USA* **56**, 1891-1898

Gilbert W, Gralla J, Mapas J, Maxam A (1975) Lactose operator sequences and the activation of *lac* repressor In: Sund H, Blauer G (Eds) *Protein-Ligand Interactions*, Walter de Gruyter, Berlin, pp 193-206

Giraldo R, Fernandez-Tresguerres E (2004) Twenty years of the pPS10 replicon: insights on the molecular mechanism for the activation of DNA replication in iteron-containing bacterial plasmids. *Plasmid* **52**, 69-83

Herrin GI, Bennet GN (1984) Role of DNA regions flanking the tryptophan promoter of *Escherichia coli* II. Insertion of *lac* operator fragments. *Gene* **32**, 349-356

Hugues TA, Pombo A, McManus J, Hozak P, Jackson DA, Cook PR (1995) On the structure of replication and transcription factories. *Journal of Cell Science Supplements* **19**, 59-65

Irani MH, Orosz L, Adhya S (1983) A control element within a structural gene: the *gal* operon of *E. coli*. *Cell* **32**, 783-788

Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**, 318-356

Kania J, Brown DT (1976) The functional repressor parts of a tetrameric *lac* repressor-β-galactosidase chimera are organized in dimers. *Proceedings of the National Academy of Sciences USA* **73**, 3529-3533

Komori H, Matsunaga F, Higuchi Y, Ishiai M, Wada C, Miki K (1999) Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6Å resolution. *The EMBO Journal* **18**, 4597- 4603

Kramer H, Niemoller M, Amouyal M, Revet B, van Wilcken-Bergmann B, Muller-Hill B (1987) *lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators. *The EMBO Journal* **6**, 1481-1487

Kramer H, Amouyal M, Nordheim A, Muller-Hill B (1988) DNA supercoiling changes the spacing requirement of two *lac* operators for DNA loop formation with *lac* repressor. *The EMBO Journal* **7**, 547-556

Krueger C, Osborne CS (2006) Raising the curtain on interchromosomal interactions. *Trends in Genetics* **22**, 637-639

Kuhnke G, Theres C, Fritz HJ, Ehring R (1989) RNA polymerase and *gal* repressor bind simultaneously and with DNA bending to the control region of the *Escherichia coli* galactose operon. *The EMBO Journal* **8**, 1247-1255

Kunnimallaiyaan S, Inman RB, Rakowski SA, Filutowicz M (2005) Role of Pi dimers in coupling ("handcuffing") of plasmid R6K's *ori* iterons. *Journal of Bacteriology* **187**, 3779-3785

Kurukuti S, Tiwari VK, Tavoosiidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R (2006) CTCF binding at the *H19* imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*. *Proceedings of the National Academy of Sciences USA* **103**, 10684-10689

Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM, Hoffman AR (2006) CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*. *Science* **312**, 269-272

McEachern M, Bott MA, Tooker PA, Helinski DR (1989) Negative control of plasmid R6K replication: possible role of intermolecular coupling of replica-tion origins. *Proceedings of the National Academy of Sciences USA* **86**, 7942-7946

Miron A, Muckerjee S, Bastia D (1992) Activation of distant replication origins *in vivo* by DNA looping as revealed by a novel mutant form of an initiator protein defective in cooperativity at a distance. *The EMBO Journal* **11**, 1205-1216

Moone H, Filippova G, Loukinov D, Pugacheva E, Chen Q, Smith ST, Munhall A, Grewe B, Bartkuhn M, Arnold R, Burke LJ, Renkawitz-Pohl R, Ohlsson R, Zhou J, Renkawitz R, Lobanenkov V (2005) CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the *Fab-8* insulator. *EMBO Reports* **6**, 165-170

Moreau P, Hen R, Wasylyk B, Everett R, Gaub MP, Chambon P (1981) The SV40 72bp repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic acids Research* **9**, 6047-6068

Mortensen L, Dandanell G, Hammer K (1989) Purification and characterization of the deoR repressor of *E. coli*. *The EMBO Journal* **8**, 325-331

Muckherjee S, Erikson H, Bastia D (1988) Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. *Cell* **52**, 375-383

Muller-Hill B, Crapo L, Gilbert W (1968) Mutants that make more *lac* repressor. *Proceedings of the National Academy of Sciences USA* **59**, 1259-1264

Murrell A, Heeson S, Reik W (2004) Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nature Genetics* **36**, 889-893

Nordstrom K (1990) Control of plasmid replication-How do iterons set the replication frequency. *Cell* **63**, 1121-1124

Oehler S, Eismann ER, Kramer H, Muller-Hill B (1990) The three operators of the *lac* operon cooperate in repression. *The EMBO Journal* **9**, 973-979

Oehler S, Amouyal M, Kolkhof P, van Wilcken-Bergmann B, Muller-Hill B (1994) Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. *The EMBO Journal* **13**, 3348-3355

O'Gorman RB, Dunaway M, Matthews KS (1980) DNA binding characteristics of *lactose* repressor and the trypsin-resistant core repressor. *The Journal of Biological Chemistry* **255**, 10100-10106

Ohlsson R, Renkawitz R, Lobanenkov V (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends in Genetics* **17**, 520-527

Pal SK, Chattoraj DK (1988) P1 plasmid replication: initiator sequestration is inadequate to explain control by initiator-binding sites. *Journal of Bacteriology* **170**, 3554-3560

Pant V, Kurukuti S, Pugacheva E, Shamsuddin S, Mariano P, Renkawitz R, Klenova E, Lobanenkov V, Ohlsssson R (2004) Mutation of a single CTCF target site within the *H19* imprinting control region leads to loss of *Igf2* imprinting and complex patterns of *de novo* methylation upon maternal inheritance. *Molecular and Cellular Biology* **24**, 3497-3504

Perez N, Rehault M, Amouyal M (2000) A functional assay in *Escherichia coli* to detect non-assisted interaction between galactose repressor dimers. *Nucleic Acids Research* **28**, 3600-3604

Perros M, Steitz TA (1996) DNA looping and *lac* repressor-CAP interaction. *Science* **274**, 1929-1930

Pfahl M, Gulde V, Bourgeois S (1979) Second and third operator of the *lac* operon: an investigation of their role in the regulatory mechanism. *Journal of Molecular Biology* **127**, 339-344

Pinkett HW, Shearwin KE, Stayrock S, Dodd IB, Burr T, Hochschild A, Egan JB, Lewis M (2006) The structural basis of cooperative regulation at an alternate genetic switch. *Molecular Cell* **2**, 605-615

Purohit PK, Nelson PC (2006) Effect of supercoiling on formation of protein-mediated DNA loops. *Physical Review E: Statistical, Non Linear and Soft Matter Physics* **74 (6 Pt 1)**, 061907

Reik W, Walter J (2001) Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics* **2**, 21-32

Reznikoff WS, Winter RB, Hurley CK (1974) The location of the repressor binding sites in the *lac* operon. *Proceedings of the National Academy of Sciences USA* **71**, 2314-2318

Sadler JR, Sasmor H, Betz JL (1983) A perfectly symmetric *lac* operator binds the *lac* repressor very tightly. *Proceedings of the National Academy of Sciences USA* **80**, 6785-6789

Schleif RF (2003) AraC protein: a love-hate relationship. *Bioessays* **25**, 274-282

Senear DF, Laue TM, Ross JB, Waxman E, Eaton S, Rusinova E (1993) The primary self-assembly of bacteriophage λcI repressor dimers is to octamer. *Biochemistry* **32**, 6179-6189

Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA (2005) Interchromosomal associations between alternatively expressed loci. *Nature* **435**, 637-645

Splinter E, Hearth H, Kooren J, Paistra RJ, Klous P, Grosveld F, Galjart N, de Laat W (2006) CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus. *Genes and Development* **20**, 2349-2354

Swan MK, Bastia D, Davies C (2006) Crystal structure of initiator-iteron complex of plasmid R6K. *Proceedings of the National Academy of Sciences USA* **103**, 18481-18486

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) Looping and interaction between hypersensitive sites in the active β-globin locus. *Molecular Cell* **10**, 1453-1465

Toukdarian AE, Helinski DR (1998) TrfA dimers play a role in copy-number control of RK2 replication. *Gene* **223**, 205-211

Twigg AJ, Sherratt D (1980) *Trans*-complementable copy-number mutants of ColE1. *Nature* **283**, 216-218

Whitson PA, Hsieh WT, Wells RD, Matthews KS (1987) Influence of DNA supercoiling and sequence context on operator DNA binding with *lac* repressor. *Journal of Biological Chemistry* **262**, 14592-14599

Winter RB, von Hippel PH (1981a) Diffusion driven mechanisms of protein translocation on nucleic acids: 2. The *E. coli lac* repressor-operator interaction: equilibrium measurements. *Biochemistry* **20**, 6948-6960

Winter RB, Berg OG, von Hippel PH (1981b) Diffusion driven mechanisms of protein translocation on nucleic acids: 3. The *E. coli lac* repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry* **20**, 6961-6977

Wu CT (1993) Transvection, nuclear structure and chromatin proteins. *Journal of Cellular Biology* **120**, 587-590

Xiong XF, de la Cruz N, Reznikoff WS (1991) Downstream deletion analysis of the *lac* promoter. *Journal of Bacteriology* **173**, 4570-4577

Xu N, Tsa CL, Lee JT (2006) Transient homologous chromosome pairing marks the onset of X inactivation. *Science* **311**, 1149-1152

Zhang Y, McEwen AE, Crothers DM, Levene SD (2006) Analysis of *in-vivo* LacR-mediated gene repression based on the mechanics of DNA looping. *PLoS ONE* **1**, e136

Zhao Z, Tavoosidddana G, Sjolinder M, Gondor A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U, Pant V, Tiwari V, Kurukuti S, Ohlsson R (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and inter-chromosomal interactions. *Nature Genetics* **38**, 1341-1347

Zzaman S, Reddy JM, Bastia D (2004) The DnaK-DnaJ-GrpE chaperone system activates inert wild type Pi initiator protein of R6K into a form active in replication initiation. *The Journal of Biological Chemistry* **279**, 50886-50894

Zzaman S, Bastia D (2005) Oligomeric initiator protein-mediated DNA looping negatively regulates plasmid replication *in vitro* by preventing origin melting. *Molecular Cell* **20**, 833-843